

Real zeroes of random polynomials, II

Descartes' rule of signs and anti-concentration on the symmetric group

Ken Söze*

Abstract

In this sequel to [8], we present a different approach to bounding the expected number of real zeroes of random polynomials with real independent identically distributed coefficients or more generally, exchangeable coefficients. We show that the mean number of real zeroes does not grow faster than the logarithm of the degree. The main ingredients of our approach are Descartes' rule of signs and a new anti-concentration inequality for the symmetric group. This paper can be read independently of part I in this series.

1 The result on the expected number of real zeroes

In this part, which can be read independently from the first part [8] (see the latter for a brief history of the problem), we will bound the expected number of real zeroes of random polynomials with real independent identically distributed coefficients or, more generally, exchangeable coefficients (the definition of exchangeability is recalled a few lines later).

For a non-zero polynomial P and a subset $A \subseteq \mathbb{R}$, let $N(A, P)$ denote the number of zeroes of P , counted with multiplicity, that fall in A . We write $N(P)$ for $N(\mathbb{R}, P)$ and $N^*(P)$ for $N(\mathbb{R} \setminus \{0\}, P)$. Everywhere in the paper, C, c, C', c' etc., denote positive numerical constants (not depending on any parameters). However, the values of these constants may change from line to line. With this notation, we are ready to state our main theorem and the key lemmas.

Theorem 1. *Let u_0, \dots, u_n , $n \geq 2$, be real numbers, not all equal to zero. Let π be a uniform random permutation of $\{0, 1, \dots, n\}$. Let $P(x) = \sum_{k=0}^n u_{\pi(k)} x^k$. Then*

$$\mathbb{E}[N^*(P)] \leq C \log n.$$

As an almost immediate corollary, we get a bound on the expected number of zeros for random polynomials with i.i.d. or, more generally, exchangeable coefficients. Recall that random variables $\lambda_1, \dots, \lambda_n$ are said to be exchangeable if the distribution of $(\lambda_{\pi(1)}, \dots, \lambda_{\pi(n)})$ is the same as the distribution of $(\lambda_1, \dots, \lambda_n)$ for any $\pi \in \mathfrak{S}_n$, where \mathfrak{S}_n is the group of permutations of $\{1, \dots, n\}$. Note that if λ_k are i.i.d., then they are exchangeable.

Corollary 2. *Let $P_n(x) = \lambda_0 + \lambda_1 x + \dots + \lambda_n x^n$, $n \geq 2$.*

1. *If $\lambda_0, \dots, \lambda_n$ are exchangeable random variables, then $\mathbb{E}[N^*(P_n) \mathbb{1}_{\{P_n \neq 0\}}] \leq C \log n$.*
2. *If $\lambda_0, \dots, \lambda_n$ are i.i.d. with $p_0 := \mathbb{P}\{\lambda_0 = 0\} < 1$, then $\mathbb{E}[N(P_n) \mathbb{1}_{\{P_n \neq 0\}}] \leq C \log n + \frac{p_0}{1-p_0}$.*

*290W 232nd Str, Apt 4b, Bronx, NY 10463, USA; sozeken65@gmail.com

The reason for the indicator function $\mathbb{1}_{\{P_n \neq 0\}}$ is that there may be a positive probability for all coefficients to vanish (in which case $N(P_n)$ is not defined). If the coefficients are i.i.d., then $\mathbb{P}\{P_n \neq 0\} = 1 - p_0^{n+1}$, showing that restricting to this event leaves out only a tiny part of the probability space, provided that n is large and p_0 is not too close to 1.

Proof. Condition on the multi-set of values $\{\lambda_0, \dots, \lambda_n\}$ (multi-set means that λ_k need not be distinct). Conditional on this multi-set being equal to $\{u_0, u_1, \dots, u_n\}$, by exchangeability, the random vector $(\lambda_0, \dots, \lambda_n)$ has the same distribution as $(u_{\pi(0)}, \dots, u_{\pi(n)})$, where π is a uniform random permutation of $\{0, 1, \dots, n\}$. On the event $P_n \neq 0$, not all u_i can equal zero, and hence Theorem 1 applies to give the first part of the corollary.

For a non-zero polynomial P , we have $N(P) = N^*(P) + N(\{0\}, P)$. By the first part, $\mathbb{E}[N^*(P_n) \mathbb{1}_{\{P_n \neq 0\}}] \leq C \log n$. If the coefficients are i.i.d., the probability that $N(\{0\}, P_n) = k$ is $p_0^k(1 - p_0)$ and hence

$$\mathbb{E}[N(\{0\}, P_n) \mathbb{1}_{\{P_n \neq 0\}}] = (1 - p_0) \sum_{k=0}^{n-1} k p_0^k \leq \frac{p_0}{1 - p_0},$$

where the last inequality follows by extending the sum up to infinity. \square

Our approach in this paper is based on Descartes' rule of signs and on the following "relative anti-concentration bound".

Lemma 3. *Let (ξ_1, \dots, ξ_k) be a vector of exchangeable random variables having a distribution that is absolutely continuous with respect to Lebesgue measure on \mathbb{R}^k . Then*

$$\mathbb{P}\left\{\left|\sum_{j=1}^k j \xi_j\right| \leq \left|\sum_{j=1}^k \xi_j\right|\right\} \leq \frac{C}{k}, \quad (1)$$

and

$$\mathbb{P}\left\{\left|\sum_{j=1}^k (-1)^j j \xi_j\right| \leq \left|\sum_{j=1}^k (-1)^j \xi_j\right|\right\} \leq \frac{C}{k}. \quad (2)$$

In turn, Lemma 3 will be deduced from an anti-concentration bound for linear forms on the symmetric group \mathfrak{S}_n . The following lemma may be considered the main technical result of this paper and potentially of interest beyond its application to proving our main theorem.

Lemma 4. *Let $n \geq 2$ and let w_1, \dots, w_n be real numbers such that $\sum_{i=1}^n w_i = 0$ and $\sum_{i=1}^n w_i^2 = 1$. Let π be a random permutation uniformly distributed on \mathfrak{S}_n . Then, for every $L \in \mathbb{R}$, we have*

$$\mathbb{P}\left\{\left|\sum_{i=1}^n w_i \pi(i) - Ln\right| \leq 1\right\} \leq \frac{C}{n} e^{-c|L|}. \quad (3)$$

We remark that it is possible to strengthen the statement to have $\frac{C}{n} e^{-cL^2}$ on the right hand side (see a brief discussion in Section 5), but we shall not need that improvement in this paper. Although the main application of this lemma is to prove Lemma 3, we shall also use it in several other smaller ways. For this purpose, we record an easy corollary of Lemma 4.

Corollary 5. *Suppose u_1, \dots, u_n are real numbers, not all equal. Let π be a random permutation uniformly distributed on \mathfrak{S}_n . Then, for every $x \in \mathbb{R}$, we have*

$$\mathbb{P}\left\{\sum_{i=1}^n u_i \pi(i) = x\right\} \leq \frac{C}{n}.$$

Proof. Set $w_i = (u_i - \bar{u})/\sigma$ where $\bar{u} = \frac{1}{n} \sum_{i=1}^n u_i$ and $\sigma^2 = \sum_{i=1}^n (u_i - \bar{u})^2$. The assumption that u_1, \dots, u_n are not all equal ensures that $\sigma > 0$ and hence w_i are well-defined, $\sum_{i=1}^n w_i = 0$ and $\sum_{i=1}^n w_i^2 = 1$. Now apply Lemma 4 with $Ln = (x - \frac{1}{2}n(n+1)\bar{u})/\sigma$. The corollary follows. \square

2 Descartes' rule of signs and deduction of Theorem 1 from Lemma 3

Recall that the number of sign-changes of a finite or infinite sequence $b = (b_0, b_1, \dots)$ of real numbers is defined as the supremum of all k for which there exist indices $0 \leq i_0 < i_1 < \dots < i_k$ such that $b_{i_j} b_{i_{j-1}} < 0$ for each $j = 1, 2, \dots, k$. Let $S(b)$ denote the number of sign changes of b .

Lemma 6 (Descartes' rule of signs [6, Chapter 5, Problem 38]). *Let $f(x) = b_0 + b_1x + b_2x^2 + \dots$ be a non-zero power series with real coefficients and convergent in $(-R, R)$. Let $N^+(f)$ be the number of zeroes (counted with multiplicity) of f in the interval $(0, R)$. Then, $N^+(f) \leq S(b)$.*

In the remaining part of this section, we prove Theorem 1 assuming that Lemma 3 and Lemma 4 are true.

2.1 Proof of Theorem 1 assuming Lemma 3 and Corollary 5

Set $P(x) = \sum_{j=0}^n \lambda_j x^j$ with $\lambda_j = u_{\pi(j)}$ where π is a uniform random permutation. Observe that $x^n P_n(1/x) = \lambda_n + \lambda_{n-1}x + \dots + \lambda_0 x^n$ is a random polynomial with the same distribution as P_n . Therefore, taking $I = (0, 1)$ and $-I = (-1, 0)$, we can write

$$\mathbb{E}[N^*(P_n)] = 2\mathbb{E}[N(I, P_n)] + 2\mathbb{E}[N(-I, P_n)] + \mathbb{E}[N(\{1\}, P_n)] + \mathbb{E}[N(\{-1\}, P_n)], \quad (4)$$

where $N(\{\pm 1\}, P_n)$ are the multiplicities of zeroes at ± 1 .

Bound for $\mathbb{E}[N(I, P_n)]$: Consider the Taylor series with radius of convergence at least 1:

$$F(x) = \frac{P_n(x)}{1-x} = \sum_{k \geq 0} S_k x^k \quad \text{and} \quad G(x) = \frac{P_n(x)}{(1-x)^2} = \sum_{k \geq 0} T_k x^k.$$

Then

$$S_k = \begin{cases} \lambda_0 + \lambda_1 + \dots + \lambda_k & \text{if } k \leq n, \\ S_n & \text{if } k > n, \end{cases}$$

and

$$T_k = \begin{cases} S_0 + S_1 + \dots + S_k = (k+1)\lambda_0 + k\lambda_1 + \dots + \lambda_k & \text{if } k \leq n, \\ T_n + (k-n)S_n & \text{if } k > n. \end{cases}$$

Firstly, $N(I, P_n) = N(I, G) \leq S((T_k)_{k \geq 0})$, by Descartes' rule. Secondly, $S((T_k)_{k \geq 0}) \leq 1 + S(T_0, T_1, \dots, T_n)$ since, beyond n , there could be at most one change of sign in the sequence $(T_k)_{k \geq n}$. Thus,

$$\mathbb{E}[N(I, P_n)] \leq 1 + \mathbb{E}[S(T_0, \dots, T_n)]. \quad (5)$$

Recall that $\lambda_j = u_{\pi(j)}$ where π is a uniform random permutation. Assume without loss of generality that on the same probability space, we have standard Gaussian random variables Z_j , $0 \leq j \leq n$, that are independent among themselves and independent of π . Set $\lambda_j^\varepsilon = \lambda_j + \varepsilon Z_j$ for $\varepsilon > 0$. Let S_k^ε and T_k^ε be defined using $(\lambda_j^\varepsilon)_{0 \leq j \leq n}$ exactly as S_k and T_k are defined in terms of $(\lambda_j)_{j \leq n}$. Then,

$$S(T_0^\varepsilon, \dots, T_n^\varepsilon) \leq \sum_{k=1}^n \mathbb{1}_{\{T_k^\varepsilon T_{k-1}^\varepsilon \leq 0\}}.$$

Since $T_k^\varepsilon = T_{k-1}^\varepsilon + S_k^\varepsilon$, to have $T_k^\varepsilon T_{k-1}^\varepsilon \leq 0$, it is necessary that $|T_k^\varepsilon| \leq |S_k^\varepsilon|$. Now, for any fixed $\varepsilon > 0$, the random variables λ_j^ε , $0 \leq j \leq n$, are exchangeable, and have an absolutely continuous distribution on \mathbb{R}^n . By conclusion (1) in Lemma 3, it immediately follows that $\mathbb{P}\{|T_k^\varepsilon| \leq |S_k^\varepsilon|\} \leq C/k$ and hence $\mathbb{E}[S(T_0^\varepsilon, \dots, T_n^\varepsilon)] \leq C \log n$. Observe that C does not depend on ε (or anything else).

Since sign changes are defined by strict inequalities, we see that almost surely,

$$S(T_0, \dots, T_n) \leq \liminf_{\varepsilon \rightarrow 0} S(T_0^\varepsilon, \dots, T_n^\varepsilon).$$

and hence, by Fatou's lemma $\mathbb{E}[S(T_0, \dots, T_n)] \leq C \log n$. Plugging back this conclusion into (5), we get $\mathbb{E}[N(I, P_n)] \leq C \log n$.

Bound for $\mathbb{E}[N(-I, P_n)]$: Next we bound $\mathbb{E}[N(-I, P_n)]$. Replacing x by $-x$, we have the analogue of (5):

$$N(-I, P_n) \leq 1 + S(T'_0, \dots, T'_n) \quad (6)$$

where $S'_k = \sum_{j=0}^k (-1)^j \lambda_j$ and $T'_k = S'_0 + \dots + S'_k = \sum_{j=0}^k (k+1-j)(-1)^j \lambda_j$.

Exactly as before, we define $\lambda_j^\varepsilon = \lambda_j + \varepsilon Z_j$, where Z_j are independent standard Gaussians that are also independent of π . Define $S_j'^\varepsilon$ and $T_j'^\varepsilon$ in terms of $(\lambda_j^\varepsilon)_{j \leq n}$ just as S'_j and T'_j are defined in terms of $(\lambda_j)_{j \leq n}$. By the lower semi-continuity of sign changes, by letting ε decrease to zero, we may deduce that $\mathbb{E}[S(T'_0, \dots, T'_n)]$ is bounded by $C \log n$ provided we prove the same bound for $\mathbb{E}[S(T_0'^\varepsilon, \dots, T_n'^\varepsilon)]$. To do that, we write

$$\begin{aligned} \mathbb{E}[S(T_0'^\varepsilon, \dots, T_n'^\varepsilon)] &\leq \sum_{k=1}^n \mathbb{P}\{T_k'^\varepsilon T_{k-1}'^\varepsilon \leq 0\} \\ &\leq \sum_{k=1}^n \mathbb{P}\{|T_k'^\varepsilon| \leq |S_k'^\varepsilon|\}. \end{aligned}$$

Now, use the bound (2) in Lemma 3 to get $\mathbb{P}\{|T_k'^\varepsilon| \leq |S_k'^\varepsilon|\} \leq C/k$. Using this bound in (6), we get the inequality $\mathbb{E}[N(-I, P_n)] \leq C \log n$.

Bound for $\mathbb{E}[N(\{\pm 1\}, P_n)]$: If $N(\{1\}, P_n) \geq 2$, we must have $P_n(1) = P'_n(1) = 0$, and therefore $\sum_{k=0}^n (k+1)u_{\pi(k)} = 0$. Obviously, this cannot happen if all u_k are equal and not zero. Then, Corollary 5 shows that this event has probability at most C/n . Therefore, $\mathbb{E}[N(\{1\}, P_n)] \leq C$ since the root at 1 has multiplicity at most n .

Now we turn to the root at -1 . If $N(\{-1\}, P_n) \geq 2$, then $P_n(-1) = P'_n(-1) = 0$ and hence, $\sum_{k=1}^{n+1} (-1)^k k \lambda_{k-1} = 0$. Using exchangeability, the probability of this event is the same as the probability of

$$\sum_{k \in E_n} \pi(k) \lambda_{k-1} = \sum_{k \in O_n} \sigma(k) \lambda_{k-1} \quad (7)$$

where π and σ are uniform random permutations of $E_n = 2\mathbb{Z} \cap \{1, 2, \dots, n+1\}$ and $O_n = (2\mathbb{Z}+1) \cap \{1, 2, \dots, n+1\}$ respectively, and π, σ are independent of each other and of $\lambda_0, \dots, \lambda_n$. Fix the values of $\lambda_0, \dots, \lambda_n$ and consider three cases.

Case 1: Suppose λ_{k-1} , $k \in E_n$ are not all equal. In this case, fix σ (i.e., condition on σ) so that the right hand side of (7) is not random anymore. We may also write $\pi(k) = 2\pi'(k/2)$ where π' is a uniform random permutation of $\frac{1}{2}E_n = \{1, 2, \dots, \lfloor \frac{n+1}{2} \rfloor\}$. Apply Corollary 5 to π' and conclude that the probability of the event in (7) is at most C/n .

Case 2: Suppose λ_{k-1} , $k \in E_n$ are all equal but λ_{k-1} , $k \in O_n$ are not all equal. Then we fix π and write $\sigma(k) = 2\sigma'((k-1)/2) + 1$ where σ' is a uniform random permutation of $\frac{1}{2}(O_n - 1) = \{0, 1, \dots, \lfloor \frac{n-1}{2} \rfloor\}$. Apply Corollary 5 to σ' and conclude that the probability of (7) is at most C/n .

Case 3: Suppose $\lambda_{k-1} = A$ for all $k \in E_n$ and $\lambda_{k-1} = B$ for all $k \in O_n$. If $A = B$, then P_n is a non-zero multiple of $1 + t + t^2 + \dots + t^n$ (recall that, by assumption, all λ_k do not vanish simultaneously) and $N(\{-1\}, P_n) \leq 1$. Hence, we assume that $A \neq B$. In this case, let τ be a uniform random permutation in $\{0, 1, \dots, n\}$ and let $\lambda'_k = \lambda_{\tau(k)}$ so that λ' has the same distribution as λ . The probability that λ'_k are equal for all $k \in E_n$ and equal for all $k \in O_n$ is smaller than e^{-cn} for some $c > 0$. Outside this event of negligible probability, λ' will fall into one of the two cases considered above.

Thus, in all cases, $\mathbb{P}\{N(\{-1\}, P_n) \geq 2\} \leq C/n$ and hence $\mathbb{E}[N(\{-1\}, P_n)] \leq C$.

In summary, we have shown that the first two terms on the right hand side of (4) are bounded by $C \log n$ and that the last two terms are bounded by C . Thus, $\mathbb{E}[N^*(P_n)] \leq C \log n$. \square

Remark: The idea of employing the sign-changes of the Taylor series of the function $(1 - x)^{-1}P_n(x)$ was used already in the pioneering paper of Bloch-Pólya [1] and then discussed by Kac [4]. Combining this idea with the classical Kolmogorov-Rogozin inequality for the concentration function (see, for instance, [2]), one can get a cruder form of Theorem 1 with \sqrt{n} in place of $\log n$.

3 Lemma 4 yields Lemma 3

The proof of Lemma 3 is based on randomization over permutations acting on (ξ_1, \dots, ξ_k) combined with estimate (3) in Lemma 4. Throughout, we say that $\pi \in \mathfrak{S}_k$ acts on the tuple $\xi = (\xi_1, \dots, \xi_k)$ by setting $(\pi\xi)_j = \xi_{\pi(j)}$; we define similarly the action of π on functions of ξ . The proof of the first estimate in Lemma 3 employs the full permutation group \mathfrak{S}_k which keeps invariant the joint distribution of the sums $\sum_{j=1}^k \xi_j$ and $\sum_{j=1}^k j\xi_j$. This is no longer possible when dealing with the sums $\sum_{j=1}^k (-1)^j \xi_j$, $\sum_{j=1}^k (-1)^j j\xi_j$, and we are forced to use subgroups of the permutation group \mathfrak{S}_k . Which subgroup to use depends on whether k is odd or even, and we distinguish between these cases in what follows.

The proof of the first estimate (1) in Lemma 3 is significantly simpler than that of the second estimate (2). The reader interested only in the case of symmetrically distributed i.i.d.s may skip the proof of (2), which is contained in Sections 3.3 and 3.4.

3.1 A corollary to Lemma 4

We start with a straightforward corollary to Lemma 4, which may be interesting on its own.

Lemma 7. *Let $n \geq 2$ and let u_1, \dots, u_n be real numbers, not all equal to zero. Let π be a uniform random permutation in \mathfrak{S}_n . Then*

$$\mathbb{P}\left\{\left|\sum_{i=1}^n u_i \pi(i)\right| \leq \left|\sum_{i=1}^n u_i\right|\right\} \leq \frac{C}{n}.$$

Proof of Lemma 7: If u_i s are all equal (and hence non-zero), then the probability in the statement is zero and there is nothing to prove. Otherwise, write $u_i = \bar{u} + \sigma w_i$, where \bar{u} is the mean of u_1, \dots, u_n , and $\sigma^2 = \sum_{i=1}^n (u_i - \bar{u})^2$. Then, w_1, \dots, w_n satisfy the hypotheses of Lemma 4. Assume without loss of generality that $\bar{u} \geq 0$. We want to get a bound on

$$\begin{aligned} \mathbb{P}\left\{\left|\sum_{i=1}^n u_i \pi(i)\right| \leq \left|\sum_{i=1}^n u_i\right|\right\} &= \mathbb{P}\left\{\left|\sum_{i=1}^n w_i \pi(i) + \frac{n(n+1)}{2} \frac{\bar{u}}{\sigma}\right| \leq n \frac{\bar{u}}{\sigma}\right\} \\ &= \mathbb{P}\left\{\sum_{i=1}^n w_i \pi(i) \in \left[\frac{n(n-1)}{2} \frac{\bar{u}}{\sigma}, \frac{n(n+3)}{2} \frac{\bar{u}}{\sigma}\right]\right\}. \end{aligned}$$

We cover the interval

$$\left[\frac{n(n-1)}{2} \frac{\bar{u}}{\sigma}, \frac{n(n+3)}{2} \frac{\bar{u}}{\sigma} \right]$$

by $\lceil \frac{n\bar{u}}{\sigma} \rceil$ intervals of length 2, and apply (3) to each subinterval (the value of L is different for different intervals, but, in any case, $|L| \geq c \frac{n\bar{u}}{\sigma}$). We get

$$\mathbb{P} \left\{ \left| \sum_{i=1}^n u_i \pi(i) \right| \leq \left| \sum_{i=1}^n u_i \right| \right\} \leq \frac{C}{n} e^{-c \frac{n\bar{u}}{\sigma}} \left\lceil \frac{n\bar{u}}{\sigma} \right\rceil. \quad (8)$$

Note that $\lceil x \rceil e^{-cx}$ is bounded by $\frac{2}{c} \vee 1$ on $[0, \infty)$ (for $x \leq 1$ the bound is 1 while for $x > 1$ we bound $\lceil x \rceil$ by $2x$ and use that $\max_{t>0} te^{-t} \leq 1$). Thus

$$\frac{C}{n} e^{-c \frac{n\bar{u}}{\sigma}} \left\lceil \frac{n\bar{u}}{\sigma} \right\rceil \leq \frac{C}{n} \left(\frac{2}{c} \vee 1 \right) \leq \frac{C'}{n}.$$

This completes the proof of Lemma 7. \square

3.2 The first estimate in Lemma 3

Here, we use Lemma 7 to deduce (1). Let A be the set $\{\xi_1, \dots, \xi_k\}$ (note that ξ_k are distinct and non-zero with probability 1). Conditional on $A = \{u_1, \dots, u_k\}$, the tuple (ξ_1, \dots, ξ_k) has the same distribution as $(u_{\pi(1)}, \dots, u_{\pi(k)})$, where π is uniformly distributed in \mathfrak{S}_k . Lemma 7 applies (as $\sum_i u_{\pi(i)} i$ has the same distribution as $\sum_i u_i \pi(i)$) to show that

$$\mathbb{P} \left\{ \left| \sum_{j=1}^k j \xi_j \right| \leq \left| \sum_{j=1}^k \xi_j \right| \mid A = \{u_1, \dots, u_k\} \right\} \leq \frac{C}{k}.$$

Since this holds for every realization of A , we get

$$\mathbb{P} \left\{ \left| \sum_{j=1}^k j \xi_j \right| \leq \left| \sum_{j=1}^k \xi_j \right| \right\} \leq \frac{C}{k},$$

completing the proof. \square

3.3 The second estimate in Lemma 3, the odd case

Let $k = 2m - 1$ and define

$$S_e = \sum_{j=1}^{m-1} \xi_{2j}, \quad S_o = \sum_{j=1}^m \xi_{2j-1}, \quad S = S_e - S_o.$$

Set $\xi'_{2j} = \xi_{2j} - \frac{1}{m-1} S_e$, $\xi'_{2j-1} = \xi_{2j-1} - \frac{1}{m} S_o$,

$$T_e = \sum_{j=1}^{m-1} j \xi'_{2j}, \quad T_o = \sum_{j=1}^m j \xi'_{2j-1}, \quad T = T_e - T_o.$$

Then

$$\sum_{j=1}^{2m-1} (-1)^j \xi_j = S_e - S_o = S,$$

and

$$\begin{aligned}
\sum_{j=1}^{2m-1} (-1)^j j \xi_j &= 2 \sum_{j=1}^{m-1} j \xi_{2j} - 2 \sum_{j=1}^m j \xi_{2j-1} + \sum_{j=1}^m \xi_{2j-1} \\
&= 2 \sum_{j=1}^{m-1} j \xi'_{2j} + 2 \frac{1}{m-1} S_e \cdot \frac{m(m-1)}{2} - 2 \sum_{j=1}^m j \xi'_{2j-1} - 2 \frac{1}{m} S_o \cdot \frac{m(m+1)}{2} + S_o \\
&= 2T_e - 2T_o + mS_e - mS_o \\
&= 2 \left[T + \frac{m}{2} S \right].
\end{aligned}$$

Thus, we aim at proving the existence of a numerical constant C so that

$$\mathbb{P} \left\{ \left| T + \frac{m}{2} S \right| \leq \frac{1}{2} |S| \right\} \leq \frac{C}{m}. \quad (9)$$

3.3.1 The subgroups \mathfrak{S}_k^e and \mathfrak{S}_k^o of \mathfrak{S}_k

Let \mathfrak{S}_k^e denote the subgroup of \mathfrak{S}_k that includes those permutations that involve only the even indices $\{2j\}_{j=1}^{m-1}$. Similarly, let \mathfrak{S}_k^o denote the subgroup of \mathfrak{S}_k that includes those permutations that involve only the odd indices $\{2j-1\}_{j=1}^m$. Let $\widehat{\mathfrak{S}}_k$ denote the subgroup of \mathfrak{S}_k consisting of permutations $\pi = \pi_e \circ \pi_o$ where $\pi_e \in \mathfrak{S}_k^e$ and $\pi_o \in \mathfrak{S}_k^o$ ($\widehat{\mathfrak{S}}_k$ is a subgroup because the subgroups \mathfrak{S}_k^e and \mathfrak{S}_k^o commute). Note that S_e and S_o , and hence S , are invariant under the action of any $\pi \in \widehat{\mathfrak{S}}_k$ on ξ ; in particular, $\pi = \pi_e \circ \pi_o \in \widehat{\mathfrak{S}}_k$ acts on $\xi' = \{\xi'_j\}_{j=1}^k$ again as a permutation. Instead of considering π drawn uniformly from \mathfrak{S}_k , we consider π_e and π_o drawn uniformly from \mathfrak{S}_k^e and \mathfrak{S}_k^o , respectively, and take $\pi = \pi_e \circ \pi_o$, writing $T_o^\pi = \pi_o T_o$, $T_e^\pi = \pi_e T_e$ and $T^\pi = T_e^\pi - T_o^\pi$. Proving (9) then reduces to proving that for any fixed tuple ξ with distinct entries,

$$\mathbb{P}^\xi \left\{ \left| T^\pi + \frac{m}{2} S \right| \leq \frac{1}{2} |S| \right\} \leq \frac{C}{m}, \quad (10)$$

where \mathbb{P}^ξ denotes averaging with respect to the product of uniform measures on \mathfrak{S}_k^e and \mathfrak{S}_k^o , with ξ fixed¹. Henceforth, we assume that $S \neq 0$, which happens almost surely because of the assumption that (ξ_1, \dots, ξ_k) has an absolutely continuous distribution.

3.3.2 Proof of estimate (2) in the case $k = 2m - 1$

Introduce the partition of \mathbb{R} determined by

$$I_j = [(j-1/2)S, (j+1/2)S), \quad j \in \mathbb{Z}.$$

¹ Here, as well as in the case $k = 2m$, we use the following observation. Let $f: \mathbb{R}^k \rightarrow \mathbb{R}$ be a Borel function (in the current instance, $f = \frac{1}{2}|S| - |T + \frac{m}{2}S|$) and let $\xi = (\xi_1, \dots, \xi_k)$ be a vector of exchangeable random variables. Then, for any subset $\tilde{\mathfrak{S}} \subset \mathfrak{S}_k$ and any probability distribution $\tilde{\mathbb{P}}$ on $\tilde{\mathfrak{S}}$, one has

$$\mathbb{P}\{f(\xi) \geq 0\} \leq \sup_{x_1, \dots, x_k} \tilde{\mathbb{P}}\{\sigma \in \tilde{\mathfrak{S}}: f(x_{\sigma_1}, \dots, x_{\sigma_k}) \geq 0\}.$$

We have then

$$\begin{aligned}
\mathbb{P}^\xi \left\{ \left| T^\pi + \frac{m}{2} S \right| \leq \frac{1}{2} |S| \right\} &= \mathbb{P}^\xi \left\{ \left| T_{\mathbf{e}}^\pi - T_{\mathbf{o}}^\pi + \frac{m}{2} S \right| \leq \frac{1}{2} |S| \right\} \\
&\leq \sum_{\substack{j_{\mathbf{e}}, j_{\mathbf{o}} \in \mathbb{Z} \\ |j_{\mathbf{e}} - j_{\mathbf{o}} + \frac{m}{2}| \leq 4}} \mathbb{P}^\xi \{ T_{\mathbf{e}}^\pi \in I_{j_{\mathbf{e}}}, T_{\mathbf{o}}^\pi \in I_{j_{\mathbf{o}}} \} \\
&= \sum_{\substack{j_{\mathbf{e}}, j_{\mathbf{o}} \in \mathbb{Z} \\ |j_{\mathbf{e}} - j_{\mathbf{o}} + \frac{m}{2}| \leq 4}} \mathbb{P}^\xi \{ T_{\mathbf{e}}^\pi \in I_{j_{\mathbf{e}}} \} \mathbb{P}^\xi \{ T_{\mathbf{o}}^\pi \in I_{j_{\mathbf{o}}} \}.
\end{aligned}$$

Note that in the range of summation of the last expression we have that $\max(|j_{\mathbf{o}}|, |j_{\mathbf{e}}|) \geq m/5$ if $m > 40$. Assuming this is the case, and using that

$$\sum_{j_{\mathbf{e}}} \mathbb{P}^\xi \{ T_{\mathbf{e}}^\pi \in I_{j_{\mathbf{e}}} \} = \sum_{j_{\mathbf{o}}} \mathbb{P}^\xi \{ T_{\mathbf{o}}^\pi \in I_{j_{\mathbf{o}}} \} = 1,$$

we obtain that

$$\mathbb{P}^\xi \left\{ \left| T^\pi + \frac{m}{2} S \right| \leq \frac{1}{2} |S| \right\} \leq \sup_{|j| \geq m/5} \mathbb{P}^\xi \{ T_{\mathbf{e}}^\pi \in I_j \} + \sup_{|j| \geq m/5} \mathbb{P}^\xi \{ T_{\mathbf{o}}^\pi \in I_j \}. \quad (11)$$

We now apply estimate (3) in Lemma 4. The argument is the same for either $T_{\mathbf{e}}$ or $T_{\mathbf{o}}$, so for concreteness set

$$\beta^2 = \sum_{j=1}^{m-1} (\xi'_{2j})^2$$

and consider the term in the right-hand side of (11) involving $T_{\mathbf{e}}$. Set $w_j = \xi'_{2j}/\beta$ so that $\sum_j w_j = 0$ and $\sum_j w_j^2 = 1$. Then, for any j ,

$$\mathbb{P}^\xi \{ T_{\mathbf{e}}^\pi \in I_j \} \leq \mathbb{P} \left\{ \left| \sum_{j=1}^{m-1} w_j \pi_{\mathbf{e}}(j) - \frac{jS}{\beta} \right| \leq \frac{|S|}{\beta} \right\}.$$

By Lemma 4, the RHS does not exceed $\frac{C}{m} e^{-\frac{cj|S|}{\beta m}} \left\lceil \frac{|S|}{\beta} \right\rceil$. Since we are interested only in j s with $|j| \geq m/5$, the latter expression is bounded by $\frac{C}{m} e^{-\frac{c|S|}{\beta}} \left\lceil \frac{|S|}{\beta} \right\rceil$. Since $\lceil x \rceil e^{-cx}$ is bounded above, we get the bound C/m .

The same argument applies with $T_{\mathbf{o}}^\pi$ replacing $T_{\mathbf{e}}^\pi$. Substituting in (11) completes the proof of the lemma when k is odd. \square

3.4 The second estimate in Lemma 3, the even case

Let $k = 2m$. We need to show that

$$\mathbb{P} \left\{ \left| \sum_{j=1}^{2m} (-1)^j j \xi_j \right| \leq \left| \sum_{j=1}^{2m} (-1)^j \xi_j \right| \right\} \leq \frac{C}{m}.$$

Put

$$S = \sum_{j=1}^{2m} (-1)^j \xi_j = \sum_{j=1}^m (\xi_{2j} - \xi_{2j-1})$$

and set $\eta_j = \xi_{2j} - \xi_{2j-1}$, noting that $S = \sum_{j=1}^m \eta_j$. As in the odd case, we center η_j by introducing $\eta'_j = \eta_j - \frac{1}{m}S$. Then

$$\begin{aligned} \sum_{j=1}^{2m} (-1)^j j \xi_j &= \sum_{j=1}^m (2j) \xi_{2j} - \sum_{j=1}^m (2j-1) \xi_{2j-1} = 2 \sum_{j=1}^m j \eta_j + \sum_{j=1}^m \xi_{2j-1} \\ &= 2 \sum_{j=1}^m j \eta'_j + (m+1)S + \sum_{j=1}^m \xi_{2j-1} = 2 \sum_{j=1}^m j \eta'_j + m\Lambda \end{aligned}$$

with

$$\Lambda = \left(1 + \frac{1}{m}\right)S + \frac{1}{m} \sum_{j=1}^m \xi_{2j-1}. \quad (12)$$

Thus, we need to estimate

$$\mathbb{P}\left\{\left|2 \sum_{j=1}^m j \eta'_j + m\Lambda\right| \leq |S|\right\}.$$

3.4.1 A randomization over local and global permutations

We introduce two subgroups of \mathfrak{S}_k . The first, which we refer to as *local* permutations, swaps the entries of the pairs (ξ_{2j-1}, ξ_{2j}) . This subgroup is generated by the transpositions τ_j , $j = 1, \dots, m$, which map $(2j-1, 2j) \rightarrow (2j, 2j-1)$. Writing $\mathfrak{S}_k^{\text{loc}}$ for the subgroup of local permutations, we note that $|\mathfrak{S}_k^{\text{loc}}| = 2^m$.

The second subgroup of \mathfrak{S}_k that we employ, which we refer to as *global* permutations, swaps the whole pairs. This subgroup is generated by the permutations $\theta_{jj'}$, $1 \leq j < j' \leq m$, which map

$$(1, 2, \dots, 2j-1, 2j, \dots, 2j'-1, 2j', \dots, k-1, k) \mapsto (1, 2, \dots, 2j'-1, 2j', \dots, 2j-1, 2j, \dots, k-1, k).$$

For instance, if we originally had the tuple $(\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6)$, we can get something like $(\xi_2, \xi_1, \xi_3, \xi_4, \xi_6, \xi_5)$ after some local permutation, and then $(\xi_3, \xi_4, \xi_6, \xi_5, \xi_2, \xi_1)$ after a global permutation. Writing $\mathfrak{S}_k^{\text{gl}}$ for the subgroup of global permutations, we note that $|\mathfrak{S}_k^{\text{gl}}| = m!$. We will consider in what follows permutations π from \mathfrak{S}_k that decompose as $\pi = \pi^{\text{g}} \circ \pi^{\text{l}}$ with $\pi^{\text{l}} \in \mathfrak{S}_k^{\text{loc}}$ and $\pi^{\text{g}} \in \mathfrak{S}_k^{\text{gl}}$ and randomize over π^{g} and π^{l} with π^{g} and π^{l} independent and uniformly distributed over $\mathfrak{S}_k^{\text{gl}}$ and $\mathfrak{S}_k^{\text{loc}}$. Note that unlike odd and even permutations considered in case k is odd, local and global permutations do not in general commute. We note that

- The quantities S and Λ are invariant under the action of $\mathfrak{S}_k^{\text{gl}}$. In what follows, the explicit form (12) of Λ will be irrelevant, only the $\mathfrak{S}_k^{\text{gl}}$ -invariance will matter.
- Random choice of π^{l} is equivalent to placing independent random signs in front of η_1, \dots, η_m .

We fix ξ (and hence, η), and denote by \mathbb{P}^ξ the law of $(\pi^{\text{g}} \circ \pi^{\text{l}})\xi$ conditioned on ξ . We need to show that

$$\mathbb{P}^\xi\left\{\left|2 \sum_{j=1}^m \pi^{\text{g}}(j)(\pi^{\text{l}}\eta)_j' + m\Lambda_{\pi^{\text{l}}}\right| \leq |S_{\pi^{\text{l}}}| \right\} \leq \frac{C}{m}.$$

Here, we put $S_{\pi^{\text{l}}} = \pi^{\text{l}}S$, $\Lambda_{\pi^{\text{l}}} = \pi^{\text{l}}\Lambda$, and, abusing notation, we denote by $\pi^{\text{g}}(j)$ the image of $j \in \{1, 2, \dots, m\}$ under π^{g} viewed as a permutation on m letters.

3.4.2 Good and bad local permutations

Put

$$\beta_{\pi^1}^2 = \sum_{j=1}^m ((\pi^1 \eta)_j')^2, \quad L_{\pi^1} = \frac{\Lambda_{\pi^1}}{2\beta_{\pi^1}}.$$

We need to estimate

$$\mathbb{P}^\xi \left\{ \left| \frac{1}{\beta_{\pi^1}} \sum_{j=1}^m \pi^g(j) (\pi^1 \eta)_j' + m L_{\pi^1} \right| \leq \frac{|S_{\pi^1}|}{2\beta_{\pi^1}} \right\}$$

We wish to apply about $|S_{\pi^1}|/\beta_{\pi^1}$ times Lemma 4 with $w_j = (\pi^1 \eta)_j'/\beta_{\pi^1}$ (this is a point where we use the randomization over $\mathfrak{S}_k^{\mathbf{g}^1}$). An obstacle is that, for some $\pi^1 \in \mathfrak{S}_k^{\text{loc}}$, the quantity β_{π^1} can be much smaller than $|S_{\pi^1}|$. The randomization over $\mathfrak{S}_k^{\text{loc}}$ will help us to circumvent this obstacle.

Put

$$B^2 = \sum_{j=1}^m \eta_j^2,$$

and note that this quantity is both $\mathfrak{S}_k^{\mathbf{g}^1}$ - and $\mathfrak{S}_k^{\text{loc}}$ -invariant. The next two claims show that outside of a tiny part of all local permutations, β_{π^1} is comparable with B .

Given a tuple $\eta = (\eta_1, \dots, \eta_m)$ we call a permutation $\pi^1 \in \mathfrak{S}_k^{\text{loc}}$ *good* if

$$\frac{1}{4}m \leq |\{j : (\pi^1 \eta)_j > 0\}| \leq \frac{3}{4}m$$

and *bad* otherwise. Let $\mathfrak{S}_k^{\text{loc,good}}$ denote the subset of $\mathfrak{S}_k^{\text{loc}}$ consisting of good permutations.

Claim 8. *For all $\pi^1 \in \mathfrak{S}_k^{\text{loc,good}}$, it holds that $\frac{1}{5}B^2 \leq \beta_{\pi^1}^2 \leq B^2$.*

Proof of Claim 8. Since

$$\beta_{\pi^1}^2 = \sum_{j=1}^m ((\pi^1 \eta)_j')^2 = \sum_{j=1}^m ((\pi^1 \eta)_j - \frac{1}{m} S_{\pi^1})^2 = \sum_{j=1}^m \eta_j^2 - \frac{1}{m} S_{\pi^1}^2 = B^2 - \frac{1}{m} S_{\pi^1}^2,$$

the upper bound $\beta_{\pi^1}^2 \leq B^2$ is immediate. It remains to prove the claimed lower bound. Note however that

$$\beta_{\pi^1}^2 = \sum_{j=1}^m ((\pi^1 \eta)_j - \frac{1}{m} S_{\pi^1})^2 \geq \sum_{j: \text{sgn}((\pi^1 \eta)_j) = -\text{sgn}(S_{\pi^1})} \left(\frac{1}{m} S_{\pi^1} \right)^2.$$

Since $\pi^1 \in \mathfrak{S}_k^{\text{loc,good}}$, the sum on the right-side is over at least $\frac{1}{4}m$ terms, and thus we obtain that

$$\beta_{\pi^1}^2 \geq \frac{1}{4} m \left(\frac{1}{m} S_{\pi^1} \right)^2 = \frac{1}{4} (B^2 - \beta_{\pi^1}^2).$$

The claim follows. \square

Thus, for good local permutations π^1 , the parameter $\beta_{\pi^1}^2$ is controlled by B^2 . The next claim asserts that most of local permutations are good:

Claim 9. *Let (ξ_1, \dots, ξ_{2m}) be an exchangeable random vector having a distribution that is absolutely continuous to Lebesgue measure on \mathbb{R}^k . Let $\eta_j = \xi_{2j} - \xi_{2j-1}$, $1 \leq j \leq m$. Then*

$$\mathbb{P}\{\pi^1 \notin \mathfrak{S}_k^{\text{loc,good}}\} \leq e^{-cm}.$$

Proof of Claim 9. Let $N_\eta^{\pi^1} = |\{j : (\pi^1 \eta)_j > 0\}|$. Note that under π^1 , the signs $\{\text{sgn}(\pi^1 \eta)_j\}_{1 \leq j \leq m}$ are i.i.d. zero mean Bernoulli random variables taking the values $\{-1, 1\}$. Letting $\{\varepsilon_i\}_{i \geq 1}$ denote i.i.d random variables taking the values $\{0, 1\}$ with equal probability, and \mathbb{P}^η denote expectation with respect to π^1 , we have that

$$\mathbb{P}^\eta \left\{ N_\eta^{\pi^1} \notin \left[\frac{1}{4}m, \frac{3}{4}m \right] \right\} \leq 2\mathbb{P} \left\{ \sum_{1 \leq i \leq m} \varepsilon_i < \frac{1}{4}m \right\} \leq e^{-cm},$$

where the last inequality follows the classical Bernstein-Hoeffding inequality (which we will recall in Section 5.3.1 below). This proves the claim. \square

3.4.3 Proof of estimate (2) in the case $k = 2m$

Given a good local permutation π^1 , we have

$$\frac{|S_{\pi^1}|}{2\beta_{\pi^1}} \leq \frac{5}{2} \frac{|S_{\pi^1}|}{B},$$

whence, by Lemma 4 applied at most $C|S_{\pi^1}|/B + 1$ times with $w_j = \pi_j^1/\beta_{\pi^1}$,

$$\mathbb{P}^{\xi, \pi^1} \left\{ \left| \frac{1}{\beta_{\pi^1}} \sum_{j=1}^m \pi^{\mathbf{g}}(j) (\pi^1 \eta)_j' + mL_{\pi^1} \right| \leq \frac{|S_{\pi^1}|}{2\beta_{\pi^1}} \right\} \leq \frac{C}{m} \left(\frac{|S_{\pi^1}|}{B} + 1 \right),$$

where \mathbb{P}^{ξ, π^1} denotes expectation with respect to $\pi^{\mathbf{g}}$.

At last, recall that, given η , $S_{\pi^1} = \sum_{j=1}^m (\pi^1 \eta)_j$ has the same distribution as $\sum_{j=1}^m \varepsilon_j \eta_j$ where ε_j are i.i.d. Bernoulli random variables taking the values $\{-1, 1\}$ with equal probability. Then, denoting by \mathbb{P}^η the expectation with respect to ε_j s, recalling that $B^2 = \sum_{j=1}^m \eta_j^2$, and using the subgaussian property of Bernoulli sums (a.k.a. the Bernstein-Chernoff inequality), we have

$$\mathbb{P}^\xi \left\{ \frac{|S_{\pi^1}|}{B} \geq t \right\} = \mathbb{P}^\eta \left\{ \left| \sum_{j=1}^m \varepsilon_j \eta_j \right| \geq tB \right\} \leq Ce^{-ct^2}, \quad t > 0,$$

whence

$$\begin{aligned} \mathbb{P}^\xi \left\{ \left| \frac{1}{\beta_{\pi^1}} \sum_{j=1}^m \pi^{\mathbf{g}}(j) (\pi^1 \eta)_j' + mL_{\pi^1} \right| \leq \frac{|S_{\pi^1}|}{2\beta_{\pi^1}}; \pi^1 \in \mathfrak{S}_k^{\text{loc, good}} \right\} \\ \leq \frac{C}{m} \left(1 + \sum_{n \geq 1} n \mathbb{P}^\xi \left\{ \frac{|S_{\pi^1}|}{B} \geq n \right\} \right) \leq \frac{C}{m}. \end{aligned}$$

Therefore, using Claim 9,

$$\mathbb{P}^\xi \left\{ \left| \frac{1}{\beta_{\pi^1}} \sum_{j=1}^m \pi^{\mathbf{g}}(j) (\pi^1 \eta)_j' + mL_{\pi^1} \right| \leq \frac{|S_{\pi^1}|}{2\beta_{\pi^1}} \right\} \leq \frac{C}{m} + e^{-cm} \leq \frac{C}{m}.$$

This completes the proof of Lemma 3 in the even case. \square

4 Anti-concentration for the symmetric group. Proof of Lemma 4

The proof of Lemma 4 goes in two steps; first, we prove the anti-concentration estimate on the length-scale \sqrt{n} :

Lemma 4'. *Let $n \geq 2$ and let w_1, \dots, w_n be real numbers such that $\sum_{i=1}^n w_i = 0$ and $\sum_{i=1}^n w_i^2 = 1$. Let π be a random permutation uniformly distributed on \mathfrak{S}_n . Then, for every $L \in \mathbb{R}$, we have*

$$\mathbb{P}\left\{\left|\sum_{i=1}^n w_i \pi(i) - Ln\right| \leq \sqrt{n}\right\} \leq \frac{C}{\sqrt{n}} e^{-c|L|}. \quad (13)$$

Then we deduce from Lemma 4' the full result, that is, the estimate on the unit-length scale. Note that Lemma 4' is weaker than Lemma 4, since the latter Lemma implies the former one.

4.1 Anti-concentration on the length-scale \sqrt{n} : proof of Lemma 4'

4.1.1 Preliminaries

Here, we introduce notation and random variables to be used in the proof of Lemma 4'.

Some notation. For each $\sigma \in \mathfrak{S}_n$, let $\Delta_\sigma = \{x \in [0, 1]^n : x_{\sigma(1)} < x_{\sigma(2)} < \dots < x_{\sigma(n)}\}$. Then Δ_σ is a simplex. The simplices Δ_σ and $\Delta_{\sigma'}$ are disjoint if $\sigma \neq \sigma'$, and the union (over all $\sigma \in \mathfrak{S}_n$) of these simplices has full Lebesgue measure in the unit cube $[0, 1]^n$. The barycenter of the simplex Δ_σ is $P_\sigma := \frac{1}{n+1} \sigma^{-1}$, where $\sigma^{-1} = (\sigma^{-1}(1), \sigma^{-1}(2), \dots, \sigma^{-1}(n))$.

As above, we denote by σu the action of the permutation $\sigma \in \mathfrak{S}_n$ on the vector $u = (u_1, \dots, u_n)$, that is, $(\sigma u)_j = u_{\sigma(j)}$. We let

$$u(\sigma)^2 = \sum_{j=1}^n (u_{\sigma(j)} + \dots + u_{\sigma(n)})^2.$$

Random variables. Next we introduce certain important random variables. Throughout, V denotes a random variable with uniform distribution in the unit cube $[0, 1]^n$. In coordinates, V_1, \dots, V_n are i.i.d. random variables with uniform distribution on $[0, 1]$. For given $\sigma \in \mathfrak{S}_n$, V_σ denotes a random variable with uniform distribution in the simplex Δ_σ . Then, $(V_\sigma)_{\sigma(1)}, \dots, (V_\sigma)_{\sigma(n)}$ have the same joint distribution as the order statistics of n independent uniform random variables on $[0, 1]$.

A bound on the density of a sum of independent uniform random variables. In the proof of Lemma 4', we will use the following lemma:

Lemma 10. *Let w_i be real (non-random) numbers and let U_i be i.i.d. random variables with uniform distribution on $[-\frac{1}{2}, \frac{1}{2}]$. Let $X = \sum_{i=1}^n w_i U_i$. Let $p_X(\cdot)$ be the density of X .*

1. *If $\sum_{i=1}^n w_i^2 = 1$, then $p_X(t) \leq C e^{-c|t|}$ for all $t \in \mathbb{R}$.*
2. *If $\sum_{i=1}^n w_i^2 \leq 1$, then $p_X(t) \leq C e^{-c|t|}$ for $|t| > 1$.*

Lemma 10 is probably known but for completeness we prove it in Section 5. We also indicate therein how to get the stronger bound $p_X(t) \leq C e^{-ct^2}$. It will become clear that if we used that improved upper bound in place of Lemma 10, we would get the bound $\frac{C}{\sqrt{n}} e^{-cL^2}$ and $\frac{C}{n} e^{-cL^2}$ in Lemmas 4' and 4, respectively.

4.1.2 Idea of the proof of Lemma 4'

Our goal is to prove estimate (13). Two difficulties are (a) discreteness of the random variable π , and (b) dependence between the random variables $\pi(1), \dots, \pi(n)$. This motivates considering first the following “baby-version” of Lemma 4 that does not have these difficulties and is a straightforward consequence of the bound for the density of $\langle w, V \rangle$ as given in Lemma 10.

A baby-version of Lemma 4'. Let V_i be i.i.d. random variables having uniform distribution on $[0, 1]$. Let w_1, \dots, w_n be real numbers satisfying $\sum_{i=1}^n w_i = 0$ and $\sum_{i=1}^n w_i^2 = 1$. Then, for any $L \in \mathbb{R}$ and $t > 0$, we have

$$\mathbb{P}\left\{\left|\sum_{i=1}^n w_i V_i - L\right| \leq t\right\} \leq Cte^{-c(|L|-t)_+}.$$

In particular, for $t = \frac{1}{\sqrt{n}}$ or $t = \frac{1}{n}$ we get the bounds $\frac{C}{\sqrt{n}}e^{-c|L|}$ and $\frac{C}{n}e^{-c|L|}$, respectively.

When we scale V_i up by n (so that they are uniform on $[0, n]$) then similarity with Lemma 4' becomes clear. Apart from analogy, observe that if π is uniformly distributed on \mathfrak{S}_n , then $\pi(i)$ (for any i) is uniformly distributed in $\{1, 2, \dots, n\}$, and any finite number of them, $\pi(i_1), \dots, \pi(i_k)$, are nearly independent (for large n).

Bad permutations. We need to count *bad* permutations π such that

$$\left|\sum_{i=1}^n w_i \pi^{-1}(i) - Ln\right| \leq \sqrt{n}.$$

This is almost the same as $|\langle w, P_\pi \rangle - L| \leq \frac{1}{\sqrt{n}}$ (not exactly the same because $P_\pi = \frac{\pi^{-1}}{n+1}$, which is not exactly $\frac{\pi^{-1}}{n}$, but that difference will be shown to be harmless). Let $\mathfrak{S}_n^{\text{bad}}$ denote the set of bad permutations.

Let $f: [0, 1]^n \rightarrow \mathbb{R}_+$ be a measurable function. Then

$$\frac{1}{n!} |\mathfrak{S}_n^{\text{bad}}| \cdot \min_{\pi \in \mathfrak{S}_n^{\text{bad}}} \mathbb{E}[f(V_\pi)] \leq \mathbb{E}[f(V)]. \quad (14)$$

The idea is to find a function f for which we can find an upper bound for $\mathbb{E}[f(V)]$ and a lower bound for $\mathbb{E}[f(V_\pi)]$ for any $\pi \in \mathfrak{S}_n^{\text{bad}}$.

A choice of the function f . The first natural choice would be $f(x) = \mathbb{1}_{\{|\langle w, x \rangle - L| \leq s\}}$ for an appropriate value of s . The reason it fails is that although $\mathbb{E}[V_\pi] = P_\pi$, there are many $\pi \in \mathfrak{S}_n^{\text{bad}}$ for which the variance of $\langle w, V_\pi \rangle$ (which is at most $w(\pi)^2/n^2$ by a simple estimate given in Claim 11 below) is quite large.

We enhance the previous choice by taking $f(x) = F(x) \mathbb{1}_{\{|\langle w, x \rangle - L| \leq s\}}$ for an appropriately chosen s , and with the choice

$$F(x) = \sum_{j=1}^n (w_{\pi(j)} + \dots + w_{\pi(n)})^2 (x_{\pi(j)} - x_{\pi(j-1)}) \quad \text{if } x \in \Delta_\pi. \quad (15)$$

The key point here is that $\mathbb{E}[F(V_\pi)] = w(\pi)^2/(n+1)$, which is large precisely for the troublesome permutations (those for which $\text{Var}[\langle w, V_\pi \rangle] \leq w(\pi)^2/n^2$ is large). Thus, we can get a better lower bound for $\mathbb{E}[f(V_\pi)]$ for $\pi \in \mathfrak{S}_n^{\text{bad}}$. It turns out that we still retain a good upper bound for $\mathbb{E}[f(V)]$.

Note that, in the proof, \mathfrak{S}_n will be broken into disjoint groups based on the value of $w(\pi)$, and inequality (14) will be applied within each group and then summed over the groups.

4.1.3 Beginning of the proof of Lemma 4'

Recall that $w(\pi)^2 = \sum_{j=1}^n (w_{\pi(j)} + \dots + w_{\pi(n)})^2$. Using Cauchy-Schwarz and the normalization $\sum_{j=1}^n w_j^2 = 1$, we see that $w(\pi)^2 \leq n^2$ for all $\pi \in \mathfrak{S}_n$. Let $Q \geq 10$ be a fixed constant (its value is unchanged throughout this section). We define the following sets whose union is all of \mathfrak{S}_n .

- $\mathfrak{S}(0) := \{\pi \in \mathfrak{S}_n : w(\pi) \leq 4(|L| + Q)\sqrt{n}\}$ and
- $\mathfrak{S}(\ell) := \{\pi \in \mathfrak{S}_n : 2^{\ell-1} < w(\pi) \leq 2^\ell\}$ for ℓ such that $4(|L| + Q)\sqrt{n} \leq 2^\ell \leq 2n$.

The goal is to prove the inequality (13). We claim that it follows if we prove that

$$\mathbb{P}\{|\langle w, P_\pi \rangle - L| \leq \frac{1}{\sqrt{n}}\} \leq \frac{C}{\sqrt{n}} e^{-c|L|}. \quad (16)$$

Indeed, what we want to bound in (13) is

$$\begin{aligned} \mathbb{P}\left\{\left|\sum_{i=1}^n w_i \pi(i) - Ln\right| \leq \sqrt{n}\right\} &= \mathbb{P}\left\{\left|\frac{1}{n+1} \sum_{i=1}^n w_i \pi(i) - L \frac{n}{n+1}\right| \leq \frac{\sqrt{n}}{n+1}\right\} \\ &\leq \mathbb{P}\left\{\left|\langle w, P_\pi \rangle - L \frac{n}{n+1}\right| \leq \frac{1}{\sqrt{n}}\right\} \end{aligned}$$

because $P_\pi = \frac{1}{n+1}\pi^{-1}$ and π^{-1} has the same distribution as π . Applying (16) we get

$$\mathbb{P}\left\{\left|\sum_{i=1}^n w_i \pi(i) - Ln\right| \leq \sqrt{n}\right\} \leq \frac{C}{\sqrt{n}} e^{-c|Ln/(n+1)|} \leq \frac{C}{\sqrt{n}} e^{-\frac{c}{2}|L|}.$$

Let $\mathfrak{S}_n^{\text{bad}} = \{\pi \in \mathfrak{S}_n : |\langle w, P_\pi \rangle - L| \leq \frac{1}{\sqrt{n}}\}$ be the set of all “bad” permutations. We shall get bounds for the cardinality of $\mathfrak{S}_n^{\text{bad}} \cap \mathfrak{S}(\ell)$ and thus get a bound for $\mathbb{P}\{\mathfrak{S}_n^{\text{bad}}\}$.

Before starting the proof, we note that it suffices to prove (16) for $n \geq n_0$ for some fixed n_0 . The reason is that for $n \leq n_0$,

$$|\langle w, P_\pi \rangle| \leq \sqrt{\sum_{i=1}^n w_i^2} \cdot \sqrt{\frac{1}{(n+1)^2} \sum_{i=1}^n (\pi^{-1}(i))^2} \leq L_0$$

for a constant L_0 (not depending on the choice of w_i s or π or n). Hence by choosing C so large that $\frac{C}{\sqrt{n_0}} e^{-c(L_0+1)} \geq 1$, the inequality (16) is trivially satisfied for all $n \leq n_0$ and for all $L \in \mathbb{R}$ (for $|L| > L_0 + 1$, the probability is zero while, for $|L| \leq L_0 + 1$, the right-hand side in (16) is bigger than 1).

4.1.4 A bound on $\text{Var}[\langle w, V_\sigma \rangle]$

To proceed, we need to bound the variance of $\langle w, V_\sigma \rangle$.

Claim 11. *Let $w = (w_1, \dots, w_n)$ be a vector in \mathbb{R}^n . Let $\sigma \in \mathfrak{S}_n$, and let V_σ be uniform on Δ_σ . Then,*

$$\text{Var}[\langle w, V_\sigma \rangle] \leq \frac{w(\sigma)^2}{(n+1)(n+2)}.$$

The proof of this claim is rather straightforward. We will give it in Section 6.

4.1.5 A bound on the cardinality of $\mathfrak{S}_n^{\text{bad}} \cap \mathfrak{S}(0)$

Let $f : [0, 1]^n \rightarrow \mathbb{R}_+$ be defined by

$$f(x) = \mathbb{1}_{\{|\langle w, x \rangle - L| \leq \frac{9(|L|+Q)}{\sqrt{n}}\}}.$$

If $\pi \in \mathfrak{S}_n^{\text{bad}} \cap \mathfrak{S}(0)$, then $w(\pi) \leq 4\sqrt{n}(|L| + Q)$ and $|\langle w, P_\pi \rangle - L| \leq \frac{1}{\sqrt{n}}$. Note that $\langle w, V_\pi \rangle$ has mean $\langle w, P_\pi \rangle$ and variance at most $\frac{1}{n^2}w(\pi)^2$ (by Claim 11). Therefore, by Chebyshev's inequality,

$$\left| \langle w, V_\pi \rangle - \langle w, P_\pi \rangle \right| \leq 2 \frac{w(\pi)}{n} \quad \text{with probability at least } \frac{1}{2}.$$

By the bound on $w(\pi)$, we see that

$$|\langle w, V_\pi \rangle - L| \leq \frac{1 + 8(|L| + Q)}{\sqrt{n}} \quad \text{with probability at least } \frac{1}{2}.$$

As $Q \geq 10$, we can write $1 + 8(|L| + Q) \leq 9(|L| + Q)$ and hence,

$$\mathbb{E}[f(V_\pi)] \geq \frac{1}{2} \quad \text{for } \pi \in \mathfrak{S}_n^{\text{bad}} \cap \mathfrak{S}(0). \quad (17)$$

Now we find an upper bound for

$$\mathbb{E}[f(V)] = \mathbb{P}\left\{|\langle w, V \rangle - L| \leq \frac{9(|L| + Q)}{\sqrt{n}}\right\}.$$

Write $\langle w, V \rangle = \sum_{i=1}^n w_i V_i = \sum_{i=1}^n w_i (V_i - \frac{1}{2})$ (since $\sum_{i=1}^n w_i = 0$) and recall that $\sum_{i=1}^n w_i^2 = 1$ to see that the first part of Lemma 10 is applicable. It gives

$$\mathbb{P}\left\{|\langle w, V \rangle - L| \leq \frac{9}{\sqrt{n}}(|L| + Q)\right\} \leq C \frac{|L| + Q}{\sqrt{n}} e^{-c(|L| - \frac{9}{\sqrt{n}}(|L| + Q))_+}. \quad (18)$$

If $|L| < 1$, we drop the exponential term, and multiply by $e^{-|L|+1}$ which is at least 1. If $|L| \geq 1$, then for $n \geq n_0$, we have $|L| - \frac{9(|L|+Q)}{\sqrt{n}} \geq \frac{1}{2}|L|$. Thus, in either case, we get the bound (for $n \geq n_0$)

$$\mathbb{E}[f(V)] \leq \frac{C}{\sqrt{n}}(|L| + Q)e^{-\frac{c}{2}|L|} \leq \frac{C}{\sqrt{n}}e^{-\frac{c}{4}|L|} \quad (19)$$

since $x \rightarrow (x + Q)e^{-cx/4}$ is bounded for $x \in (0, \infty)$.

Invoking (14), we conclude from (17) and (19) that, for $n \geq n_0$ and for all $L \in \mathbb{R}$,

$$\frac{1}{n!} |\mathfrak{S}_n^{\text{bad}} \cap \mathfrak{S}(0)| \leq \frac{C}{\sqrt{n}} e^{-c|L|}. \quad (20)$$

4.1.6 A bound on the cardinality of $\mathfrak{S}_n^{\text{bad}} \cap \mathfrak{S}(\ell)$ with $4(|L| + Q)\sqrt{n} \leq 2^\ell \leq 2n$

Fix $T = 2^\ell$ so that $\frac{T}{2} \leq w(\pi) \leq T$. Let $f_T : [0, 1]^n \rightarrow \mathbb{R}_+$ be defined as

$$f_T(x) = F(x) \mathbb{1}_{\{|\langle w, x \rangle - L| \leq \frac{2QT}{n}\}},$$

where F is the function that was defined in (15).

A lower bound for $\mathbb{E}[f(V_\pi)]$, $\pi \in \mathfrak{S}_n^{\text{bad}} \cap \mathfrak{S}(\ell)$. The random variable $\langle w, V_\pi \rangle$ has mean $\langle w, P_\pi \rangle$ and Claim 11 asserts that $\text{Var}[\langle w, V_\pi \rangle] \leq \frac{T^2}{n^2}$. By Chebyshev's inequality,

$$|\langle w, V_\pi \rangle - \langle w, P_\pi \rangle| \leq Q \frac{T}{n} \quad \text{with probability at least } 1 - \frac{1}{Q^2}.$$

If $\pi \in \mathfrak{S}_n^{\text{bad}} \cap \mathfrak{S}(\ell)$, then, in addition to the above, we have $|\langle w, P_\pi \rangle - L| \leq \frac{1}{\sqrt{n}}$. Therefore, recalling that $Q \geq 10$,

$$|\langle w, V_\pi \rangle - L| \leq \frac{\sqrt{n} + QT}{n} \quad \text{with probability at least } 0.99. \quad (21)$$

By the definition (15), we can write

$$F(V_\pi) = \sum_{j=1}^n \alpha_j ((V_\pi)_{\pi(j)} - (V_\pi)_{\pi(j-1)}) \quad \text{with } \alpha_j = (w_{\pi(j)} + \dots + w_{\pi(n)})^2.$$

Then, recall that $\mathbb{E}[V_\pi] = P_\pi$ which is $\frac{1}{n+1}\pi^{-1}$. Hence, $\mathbb{E}[(V_\pi)_{\pi(j)}] = \frac{j}{n+1}$ and

$$\mathbb{E}[F(V_\pi)] = \frac{1}{n+1} \sum_{j=1}^n \alpha_j = \frac{1}{n+1} w(\pi)^2.$$

We may also rewrite $F(V_\pi)$ as $\sum_{j=1}^n (\alpha_j - \alpha_{j+1})(V_\pi)_{\pi(j)}$ (with the convention that $\alpha_{n+1} = 0$). Then, applying Claim 11 we get $\text{Var}[F(V_\pi)] \leq \frac{1}{(n+1)^2} \sum_{j=1}^n \alpha_j^2$. By the second moment inequality²

$$\begin{aligned} \mathbb{P}\{F(V_\pi) \geq \frac{1}{2} \mathbb{E}[F(V_\pi)]\} &\geq \frac{1}{4} \frac{(\mathbb{E}[F(V_\pi)])^2}{\text{Var}[F(V_\pi)] + (\mathbb{E}[F(V_\pi)])^2} \\ &= \frac{1}{4} \left(1 + \frac{\text{Var}[F(V_\pi)]}{(\mathbb{E}[F(V_\pi)])^2}\right)^{-1} \\ &\geq \frac{1}{4} \left(1 + \frac{\sum_{j=1}^n \alpha_j^2}{(\sum_{j=1}^n \alpha_j)^2}\right)^{-1}. \end{aligned}$$

Since α_j s are all non-negative, $(\sum_j \alpha_j)^2 \geq \sum_j \alpha_j^2$ and hence

$$\mathbb{P}\{F(V_\pi) \geq \frac{1}{2} \mathbb{E}[F(V_\pi)]\} \geq \frac{1}{8}.$$

Combine this with (21) to conclude that

$$\mathbb{P}\{F(V_\pi) \geq \frac{1}{2(n+1)} w(\pi)^2 \text{ and } |\langle w, V_\pi \rangle - L| \leq \frac{2QT}{n}\} \geq \frac{1}{10}.$$

Consequently,

$$\mathbb{E}[f_T(V_\pi)] \geq \frac{w(\pi)^2}{2(n+1)} \cdot \frac{1}{10} \stackrel{w(\pi) \geq T/2}{\geq} \frac{T^2}{80(n+1)} \geq \frac{T^2}{160n} \quad \text{for } \pi \in \mathfrak{S}_n^{\text{bad}} \cap \mathfrak{S}(\ell). \quad (22)$$

² The second moment inequality asserts that if X is a non-negative random variable, then $\mathbb{P}\{X \geq \lambda \mathbb{E}[X]\} \geq (1 - \lambda)^2 \frac{(\mathbb{E}[X])^2}{\mathbb{E}[X^2]}$ for $0 < \lambda < 1$. The proof is a straightforward application of the Cauchy-Schwarz inequality, see [5, p.8].

An upper bound for $\mathbb{E}[f_T(V)]$. For this, we write $F(x)$ in the following alternative form.

$$F(x) = \int_0^1 (G_t(x))^2 dt \quad \text{where } G_t(x) = \sum_{j=1}^n w_j \mathbb{1}_{\{x_j > t\}}.$$

It is easy to check that this agrees with (15). Hence,

$$\mathbb{E}[f_T(V)] = \int_0^1 \mathbb{E}[G_t(V)^2 \mathbb{1}_{\{|\langle w, V \rangle - L| \leq \frac{2QT}{n}\}}] dt. \quad (23)$$

Fix $t \in (0, 1)$ and let $S_t = \{i: V_i > t\}$. We estimate the integrand for each t .

Write

$$\begin{aligned} \mathbb{E}\left[G_t(V)^2 \mathbb{1}_{\{|\langle w, V \rangle - L| \leq \frac{2QT}{n}\}}\right] &\leq 64(|L| + 2Q)^2 \mathbb{P}\left\{|\langle w, V \rangle - L| \leq \frac{2QT}{n}\right\} \\ &\quad + \mathbb{E}\left[G_t(V)^2 \mathbb{1}_{\{|G_t(V)| \geq 8(|L| + 2Q)\}} \mathbb{1}_{\{|\langle w, V \rangle - L| \leq \frac{2QT}{n}\}}\right]. \end{aligned} \quad (24)$$

We bound the first term along similar lines to the case $\ell = 0$. Since $\sum_{i=1}^n w_i^2 = 1$, we may apply the first part of Lemma 10 to the random variable $\langle w, V \rangle$ to get

$$\mathbb{P}\left\{|\langle w, V \rangle - L| \leq \frac{2QT}{n}\right\} \leq C \frac{2QT}{n} \exp\left\{-c\left(|L| - \frac{2QT}{n}\right)_+\right\}. \quad (25)$$

Since $T \leq 2n$ we see that $\frac{2QT}{n} \leq 4Q$, a constant. Hence, dividing into the cases $|L| \geq 8Q$ and $|L| \leq 8Q$, and changing constants suitably, for $n \geq n_0$ and for all $L \in \mathbb{R}$ we have the inequality

$$\mathbb{P}\left\{|\langle w, V \rangle - L| \leq \frac{2QT}{n}\right\} \leq C \frac{T}{n} \exp\{-c|L|\}.$$

Thus, the first term in (24) is bounded by (for $n \geq n_0$ and for all $L \in \mathbb{R}$)

$$64(|L| + 2Q)^2 \mathbb{P}\left\{|\langle w, V \rangle - L| \leq \frac{2QT}{n}\right\} \leq C \frac{T}{n} (|L| + 2Q)^2 e^{-c|L|} \leq C \frac{T}{n} e^{-c|L|/2}, \quad (26)$$

again because $x \mapsto (x + 2Q)^2 e^{-cx/2}$ is bounded.

It remains to control the second term in (24). The trick is to condition on the random set $S_t = \{i: V_i > t\}$. We need to understand the conditional distributions of the two random variables $G_t(V)$ and $\langle w, V \rangle$. The first one is easy since $G_t(V) = \sum_{i \in S_t} w_i$, which is a function of S_t . In other words, conditional on S_t , the variable $G_t(V)$ is a constant.

Next, conditional on the set S_t , the random variables V_1, \dots, V_n are still independent, V_i is uniformly distributed on $[t, 1]$ if $i \in S_t$ and V_i is uniformly distributed in $[0, t]$ if $i \notin S_t$. Let U_i be independent random variables distributed uniformly on $[-\frac{1}{2}, \frac{1}{2}]$ and set

$$V'_i = \begin{cases} \frac{1}{2}t + tU_i & \text{if } i \notin S_t, \\ \frac{1}{2}(1+t) + (1-t)U_i & \text{if } i \in S_t. \end{cases}$$

Then, the distribution of the vector $V' = (V'_1, \dots, V'_n)$ is the same as the conditional distribution of V given S_t . In particular, the conditional distribution of $\langle w, V \rangle$ given S_t is the same as the unconditional distribution of

$$\begin{aligned} \langle w, V' \rangle &= \frac{1}{2}(1+t) \sum_{i \in S_t} w_i + \sum_{i \in S_t} w_i(1-t)U_i + \frac{1}{2}t \sum_{i \notin S_t} w_i + \sum_{i \notin S_t} w_i t U_i \\ &= \frac{1}{2} G_t(V) + \sum_{i \in S_t} w_i(1-t)U_i + \sum_{i \notin S_t} w_i t U_i. \end{aligned}$$

The quantity we need to control is $\mathbb{E} \left[G_t(V)^2 \mathbb{1}_{\{|G_t(V)| \geq 8(|L|+2Q)\}} \mathbb{1}_{\{|\langle w, V \rangle - L| \leq \frac{2QT}{n}\}} \right]$. By conditioning on S_t , we may write this as

$$\mathbb{E} \left[G_t(V)^2 \mathbb{1}_{\{|G_t(V)| \geq 8(|L|+2Q)\}} \mathbb{E} \left[\mathbb{1}_{\{|\langle w, V \rangle - L| \leq \frac{2QT}{n}\}} \mid S_t \right] \right]. \quad (27)$$

(As $G_t(V)$ is a function of S_t , factors involving it can be taken out of the conditional expectation.) Using our representation of the conditional distribution of $\langle w, V \rangle$ in terms of the U_i s, the inner conditional expectation may be written as

$$\begin{aligned} \mathbb{E} \left[\mathbb{1}_{\{|\langle w, V \rangle - L| \leq \frac{2QT}{n}\}} \mid S_t \right] &= \mathbb{E} \left[\mathbb{1}_{\{|\langle w', V \rangle - L| \leq \frac{2QT}{n}\}} \mid S_t \right] \\ &= \mathbb{P} \left\{ \left| \sum_{i=1}^n w'_i U_i + \frac{1}{2} G_t(V) - L \right| \leq \frac{2QT}{n} \mid S_t \right\}, \end{aligned}$$

where $w'_i = (1-t)w_i$ if $i \in S_t$ and $w'_i = tw_i$ if $i \notin S_t$. Again, we want to apply the density bound from Lemma 10. However, we only have the upper bound $\sum_{i=1}^n (w'_i)^2 \leq 1$, and hence, to apply the second part of that lemma, we need to make sure that at least on the event $\{|G_t(V)| \geq 8(|L| + 2Q)\}$, see (27), the interval

$$\left[\frac{1}{2} G_t(V) - L - \frac{2QT}{n}, \frac{1}{2} G_t(V) - L + \frac{2QT}{n} \right]$$

is at distance at least 1 from the origin. Once we show that, the second part of Lemma 10 gives the bound

$$\mathbb{E} \left[\mathbb{1}_{\{|\langle w, V \rangle - L| \leq \frac{2QT}{n}\}} \mid S_t \right] \leq C \frac{2QT}{n} \exp \left\{ -c \left(\frac{1}{2} |G_t(V)| - |L| - \frac{2QT}{n} \right)_+ \right\}. \quad (28)$$

But on the event $\{|G_t(V)| \geq 8(|L| + 2Q)\}$, recalling that $\frac{2QT}{n} \leq 4Q$, one has

$$\left| \frac{1}{2} G_t(V) - L \pm \frac{2QT}{n} \right| \geq \frac{1}{2} |G_t(V)| - |L| - 4Q \geq 3|L| + 4Q,$$

which is at least 1. Thus, the bound (28) is valid.

We now multiply the left hand side of (28) by $G_t(V)^2 \mathbb{1}_{\{|G_t(V)| \geq 8(|L|+2Q)\}}$ and take expectations. If the indicator is to be non-zero, then

$$\frac{1}{2} G_t(V) - |L| - \frac{2QT}{n} \geq \frac{1}{8} G_t(V) + 3(|L| + 2Q) - |L| - 4Q = \frac{1}{8} G_t(V) + 2|L| + 2Q.$$

Using this bound along with (28), we get

$$\text{the second term in (24)} \leq C \frac{T}{n} \mathbb{E} \left[G_t(V)^2 \exp \left\{ -c \left(\frac{1}{8} G_t(V) + 2|L| + 2Q \right) \right\} \right] \leq C \frac{T}{n} e^{-c|L|}$$

since $G_t(V)^2 \exp\{-cG_t(V)\}$ is bounded by a constant.

Adding this to the bound for the first term given in (24), we arrive at

$$\mathbb{E} \left[G_t(V)^2 \mathbb{1}_{\{|\langle w, V \rangle - L| \leq \frac{2QT}{n}\}} \right] \leq C \frac{T}{\sqrt{n}} e^{-c|L|}.$$

Integrating this bound over t and plugging into (23) gives us

$$\mathbb{E}[f_T(V)] \leq C \frac{T}{n} e^{-c|L|}. \quad (29)$$

Tying the ends together. At last, juxtaposing (22) and (29), and recalling that $T = 2^\ell$, we conclude that

$$\frac{1}{n!} |\mathfrak{S}_n^{\text{bad}} \cap \mathfrak{S}(\ell)| \leq \frac{\mathbb{E}[f_T(V)]}{\min_{\pi \in \mathfrak{S}_n^{\text{bad}}} \mathbb{E}[f_T(V_\pi)]} \leq \frac{Cn}{T^2} \cdot \frac{CT}{n} e^{-c|L|} = \frac{C}{T} e^{-c|L|} = \frac{C}{2^\ell} e^{-c|L|}. \quad (30)$$

4.1.7 End of the proof of Lemma 4'

Adding up (30) over ℓ such that $2^\ell \geq 4(|L| + Q)\sqrt{n}$ and recalling the estimate for $\frac{1}{n!} |\mathfrak{S}_n^{\text{bad}} \cap \mathfrak{S}(0)|$ we got in (20), we obtain

$$\begin{aligned} \frac{1}{n!} |\mathfrak{S}_n^{\text{bad}}| &\leq \frac{1}{n!} |\mathfrak{S}_n^{\text{bad}} \cap \mathfrak{S}(0)| + \sum_{2^\ell \geq 4(|L| + Q)\sqrt{n}} \frac{1}{n!} |\mathfrak{S}_n^{\text{bad}} \cap \mathfrak{S}(\ell)| \\ &\leq \frac{C}{\sqrt{n}} e^{-c|L|} + \sum_{2^\ell \geq 4(|L| + Q)\sqrt{n}} \frac{C}{2^\ell} e^{-c|L|} \leq \frac{C}{\sqrt{n}} e^{-c|L|}. \end{aligned}$$

This completes the proof of Lemma 4'. \square

4.2 From the scale \sqrt{n} to the unit length-scale

For the permutation $\pi \in \mathfrak{S}_n$, define its “weight” as $\text{wt}[\pi] = \sum_{j=1}^n \pi(j)w_j = \sum_{j=1}^n jw_{\pi(j)}$. Fix L and define the set of “bad permutations” $\mathfrak{S}_n^{\text{bad}}$ as the set of all $\pi \in \mathfrak{S}_n$ for which $|\text{wt}[\pi] - Ln| \leq 1$.

4.2.1 Partition of \mathfrak{S}_n

For each permutation $\sigma = (\sigma(1), \dots, \sigma(n-1)) \in \mathfrak{S}_{n-1}$, let

$$\begin{aligned} \widehat{\mathfrak{S}}_\sigma &= \{\pi_\sigma^{(1)} = (n, \sigma(1), \dots, \sigma(n-1)), \pi_\sigma^{(2)} = (\sigma(1), n, \sigma(2), \dots, \sigma(n-1)), \dots \\ &\dots, \pi_\sigma^{(n)} = (\sigma(1), \dots, \sigma(n-1), n)\}. \end{aligned}$$

Then, $\{\widehat{\mathfrak{S}}_\sigma : \sigma \in \mathfrak{S}_{n-1}\}$ is a partition of \mathfrak{S}_n into groups of n permutations each. The key point is to show that, for most σ , there are only a few bad permutations in $\widehat{\mathfrak{S}}_\sigma$.

4.2.2 The subset $\mathcal{A} \subset \mathfrak{S}_{n-1}$

Let $b = \max_j w_j - \min_j w_j$ and note that, due to normalization $\sum w_j = 0$ and $\sum w_j^2 = 1$, we have $\frac{1}{\sqrt{n}} \leq b \leq 2$. We will need two subsets, \mathcal{A} and \mathcal{E} of \mathfrak{S}_{n-1} . These sets are “exceptional”, in the sense that they have small cardinality. Here, we will define \mathcal{A} , the set \mathcal{E} will be defined later.

Let \mathcal{A} be the set of $\sigma \in \mathfrak{S}_{n-1}$ for which $\text{wt}[\pi] \in [Ln - 1 - nb, Ln + 1 + nb]$ for all $\pi \in \widehat{\mathfrak{S}}_\sigma$. Then,

$$\begin{aligned} \frac{1}{(n-1)!} |\mathcal{A}| &\leq \frac{1}{n!} \left| \{\pi \in \mathfrak{S}_n : \text{wt}[\pi] \in [Ln - 1 - nb, Ln + 1 + nb]\} \right| \\ &= \mathbb{P}\{\text{wt}[\pi] \in [Ln - 1 - nb, Ln + 1 + nb]\}. \end{aligned}$$

Divide the interval $[Ln - 1 - nb, Ln + 1 + nb]$ into $\lceil \frac{nb+1}{\sqrt{n}} \rceil$ intervals of length $2\sqrt{n}$ (or less) each, and invoke Lemma 4' for each of them. All the intervals are at distance at least $(|L| - b - \frac{1}{n})n$ from the origin, and hence

$$\frac{1}{(n-1)!} |\mathcal{A}| \leq \left\lceil \frac{nb+1}{\sqrt{n}} \right\rceil \frac{C}{\sqrt{n}} e^{-c(|L|-3)_+} \leq Cbe^{-c|L|}. \quad (31)$$

4.2.3 A relabelling of w_1, \dots, w_n

To define the exceptional set \mathcal{E} we consider three cases: $b \geq \frac{1}{12}$, $n^{-\frac{1}{3}} \leq b \leq \frac{1}{12}$, and $n^{-\frac{1}{2}} \leq b \leq n^{-\frac{1}{3}}$. First, we arrange the w_i s in a suitable manner:

1. $b \geq \frac{1}{12}$. Label w_i s so that $w_n = \max_j w_j$ or $w_n = \min_j w_j$, whichever of the two is larger in absolute value.
2. $n^{-\frac{1}{3}} \leq b \leq \frac{1}{12}$. As in the first case, we label w_i s so that $w_n = \max_j w_j$ or $w_n = \min_j w_j$, whichever of the two is larger in absolute value.
3. $\frac{1}{\sqrt{n}} \leq b \leq n^{-\frac{1}{3}}$. Label w_i s so that $w_n = \max_j w_j$ or $w_n = \min_j w_j$,

whichever of the two is such that the cardinality of $\{j: |w_n - w_j| \geq \frac{b}{2}\}$ is at least $\frac{n}{2}$.

For simplicity of language, let us assume that $w_n = \max_j w_j$ in all cases (in fact, there is no loss of generality as we may negate all the w_j s and L if needed). Then, for any $\sigma \in \mathfrak{S}_{n-1}$ and any $k \leq n-1$, we have

$$\text{wt}[\pi_\sigma^{(k+1)}] - \text{wt}[\pi_\sigma^{(k)}] = w_n - w_{\sigma(k)}. \quad (32)$$

As $w_n = \max_j w_j$, these increments are non-negative. Thus, for any σ ,

- we have $\text{wt}[\pi_\sigma^{(n)}] - \text{wt}[\pi_\sigma^{(1)}] \leq nb$, and therefore, if $\sigma \notin \mathcal{A}$, then $\mathfrak{S}_n^{\text{bad}} \cap \widehat{\mathfrak{S}}_\sigma$ is empty;
- the set of k for which $\pi_\sigma^{(k)}$ is bad, is a discrete interval of the form $\{k_\sigma, k_\sigma+1, \dots, k_\sigma+\ell_\sigma-1\}$.

4.2.4 The exceptional set $\mathcal{E} \subset \mathfrak{S}_{n-1}$

In each of the three cases, we define the exceptional set \mathcal{E} and get an upper bound for ℓ_σ for $\sigma \notin \mathcal{E}$. We also need to control the cardinality of \mathcal{E} , of course.

1st case: $b \geq \frac{1}{12}$. In this case $w_n \geq \frac{b}{2}$. Let $A = \{k: w_k \geq \frac{b}{4}\}$ and observe that $|A| \leq \frac{16}{b^2}$ since $\sum w_i^2 = 1$.

Now fix any $\sigma \in \mathfrak{S}_{n-1}$. If, for some $k \in \{1, 2, \dots, n-1\}$, $\sigma_k \notin A$, then, by (32), we see that $\text{wt}[\pi_\sigma^{(k+1)}] - \text{wt}[\pi_\sigma^{(k)}] > \frac{b}{4}$. Recall that bad permutations are those for which $\text{wt}[\pi]$ is in $[Ln-1, Ln+1]$, an interval of length 2. This shows that

$$\left(\ell_\sigma - 1 - \frac{16}{b^2}\right) \frac{b}{4} \leq w_{k_\sigma+\ell_\sigma-1} - w_{k_\sigma} \leq 2.$$

Hence $\ell_\sigma \leq \frac{16}{b^2} + \frac{8}{b} + 1$ for all $\sigma \in \mathfrak{S}_{n-1}$. Set $C = \frac{16}{(1/12)^2} + \frac{8}{(1/12)} + 1$. Let $\mathcal{E} = \emptyset$ (no need in exceptional permutations in this case). Then,

$$|\mathcal{E}| = 0 \text{ and } \ell_\sigma \leq C \text{ for all } \sigma \in \mathfrak{S}_n. \quad (33)$$

2nd case: $n^{-\frac{1}{3}} \leq b \leq \frac{1}{12}$. Fix σ and let $I_\sigma = \{\sigma(j): k_\sigma \leq j < k_\sigma + \ell_\sigma - 1\}$ and $A = \{k: w_k \geq \frac{b}{4}\}$. Exactly as in the previous case, $|A| \leq \frac{16}{b^2}$ (which may be rather large in the present case, hence further arguments below) and

$$(\ell_\sigma - 1 - |A \cap I_\sigma|) \frac{b}{4} \leq 2.$$

Therefore, if $\ell_\sigma \geq \frac{10}{b}$, then $|A \cap \{\sigma(j): k_\sigma \leq j < k_\sigma + \frac{10}{b} - 1\}| \geq \frac{1}{b}$. Let $m = \lceil \frac{1}{b} \rceil$ and define

$$\mathcal{E} = \{\sigma: |A \cap \{\sigma(j): k \leq j \leq k + 10m - 1\}| \geq m \text{ for some } k\}.$$

Then, $\ell_\sigma \leq \frac{10}{b}$ for $\sigma \notin \mathcal{E}$.

We now need to bound the cardinality of the exceptional set \mathcal{E} . For any $\sigma \in \mathcal{E}$, we associate a $(2m+1)$ -tuple $(k, j_1, \dots, j_m, a_1, \dots, a_m)$, where $1 \leq k \leq j_1 < \dots < j_m < k + 10m - 1 \leq n$ and a_1, \dots, a_m are elements of A such that $\sigma(j_1) = a_1, \dots, \sigma(j_m) = a_m$. By definition of \mathcal{E} , such a tuple exists, and if there is more than one, make an arbitrary choice to fix one.

Now we get an upper bound for the number of distinct $(2m+1)$ -tuples that can arise in this manner. Firstly, the number of choices for k is less than n , and having chosen k , the numbers j_1, \dots, j_m may be chosen in less than 2^{10m} ways (any subset of $\{k, k+1, \dots, k+10m-1\}$) and after that choice, a_1, \dots, a_m may be chosen in at most $|A|^m$ ways.

Finally, any given $(2m+1)$ -tuple can come from at most $(n-m)!$ permutations, since the values of $\sigma(j_1), \dots, \sigma(j_m)$ are fixed. Thus,

$$\frac{1}{(n-1)!} |\mathcal{E}| \leq \frac{(n-m)!}{(n-1)!} n 2^{10m} \left(\frac{16}{b^2}\right)^m \leq \frac{n}{(n/2)^{\frac{1}{b}-1}} \left(\frac{C}{b^2}\right)^{1/b} \leq n^2 \left(\frac{C}{nb^2}\right)^{1/b}$$

since $|A| \leq \frac{16}{b^2}$ and $n-m \geq \frac{n}{2}$ (as $m = \lceil \frac{1}{b} \rceil \leq 2\sqrt{n}$). If $b \geq n^{-1/3}$ (in fact we may go up to $b \geq n^{-\frac{1}{2}+\delta}$ for any $\delta > 0$), it is easy to see that the above quantity is bounded by

$$n^2 \left(\frac{C}{n^{-2/3} \cdot n}\right)^{12} \leq \frac{C}{n^2}.$$

These manipulations are all valid for $n \geq n_0$ for some fixed n_0 . Thus,

$$\frac{1}{(n-1)!} |\mathcal{E}| \leq \frac{C}{n^2} \quad \text{and} \quad \ell_\sigma \leq \frac{C}{b} \quad \text{for all } \sigma \notin \mathcal{E}, \quad (34)$$

where the constant C can take care of all the cases when $n < n_0$.

3rd case: $\frac{1}{\sqrt{n}} \leq b \leq n^{-\frac{1}{3}}$. Setting $A = \{k: w_n - w_k < \frac{b}{2}\}$ we see that $|A| \leq \frac{n}{2}$ (because of the way we chose w_n). Fix σ and let $I_\sigma = \{\sigma(j): k_\sigma \leq j < k_\sigma + \ell_\sigma - 1\}$. Again

$$(\ell_\sigma - 1 - |A \cap I_\sigma|) \frac{b}{2} \leq 2.$$

Hence if $\ell_\sigma \geq \frac{12}{b}$, then $|A \cap I_\sigma| \geq \frac{7}{b}$. Define

$$\mathcal{E} = \left\{ \sigma: |A \cap \{\sigma(j): k \leq j < k + \frac{12}{b}\}| \geq \frac{7}{b} \text{ for some } k \right\}.$$

Then, $\ell_\sigma \leq \frac{12}{b}$ for $\sigma \notin \mathcal{E}$. We want to bound the cardinality of \mathcal{E} .

Let $m = \lceil \frac{1}{b} \rceil$ and let $\mathcal{E} = \bigcup_k \mathcal{E}_k$ where

$$\mathcal{E}_k = \left\{ \sigma: |A \cap \{\sigma(j): k \leq j \leq k + 12m - 1\}| \geq 7m \right\}.$$

Fix $k \leq n - 12m + 1$. The quantity $\frac{1}{(n-1)!} |\mathcal{E}_k|$ has the following interpretation. Suppose we have a basket with $n-1$ different balls labeled by $\{1, 2, \dots, n-1\}$, $|A| \leq \frac{1}{2}n$ of these balls are black, while the rest are white. We take, at random, $12m$ different balls (without returning them to the basket). Then, $\frac{1}{(n-1)!} |\mathcal{E}_k|$ is the probability that at least $7m$ of these $12m$ balls will be black. Whence, using e.g. Stirling's formula, for $n \geq n_0$,

$$\frac{1}{(n-1)!} |\mathcal{E}_k| \leq e^{-cm}.$$

This is the bound for fixed k . Add over k to see that $\frac{1}{(n-1)!} |\mathcal{E}| \leq ne^{-cm}$. Since $m \geq n^{1/3}$, this gives

$$\frac{1}{(n-1)!} |\mathcal{E}| \leq \frac{C}{n^2} \text{ and } \ell_\sigma \leq \frac{C}{b} \text{ for all } \sigma \notin \mathcal{E}. \quad (35)$$

Again, the constant C is adjusted so that the above estimates are also valid for $n < n_0$. This completes the third case.

4.2.5 Completing the proof of Lemma 4

To finish the proof, recall that if $\sigma \notin \mathcal{A}$, then $\mathfrak{S}_n^{\text{bad}} \cap \widehat{\mathfrak{S}}_\sigma$ is empty. Consequently,

$$\frac{1}{n!} |\mathfrak{S}_n^{\text{bad}}| \leq \frac{1}{n!} \sum_{\sigma \in \mathcal{A}} \ell_\sigma \leq \frac{1}{n!} \frac{C}{b} |\mathcal{A}| + \frac{1}{n!} n |\mathcal{A} \cap \mathcal{E}|,$$

by applying the bound $\ell_\sigma \leq \frac{C}{b}$ for $\sigma \notin \mathcal{E}$ as given in (33), (34) and (35) and the trivial bound $\ell_\sigma \leq n$ for $\sigma \in \mathcal{E}$. Note that $|\mathcal{A} \cap \mathcal{E}| \leq |\mathcal{A}| \wedge |\mathcal{E}| \leq \sqrt{|\mathcal{A}|} \sqrt{|\mathcal{E}|}$. From the bounds (33), (34) and (35) on the cardinality of \mathcal{E} and the bound (31) on the cardinality of \mathcal{A} , we get

$$\frac{1}{n!} |\mathfrak{S}_n^{\text{bad}}| \leq \frac{C}{n} e^{-c|L|} + \frac{1}{(n-1)!} \sqrt{(n-1)! \frac{C}{n^2}} \cdot \sqrt{(n-1)! C b e^{-c|L|}} \leq \frac{C'}{n} e^{-c'|L|}.$$

This completes the proof of Lemma 4. \square

5 The proof of Lemma 10

We conclude the paper by proving Lemma 10, which says that the density of the distribution of the sum $X = \sum_{i=1}^n w_i U_i$ (where U_i are i.i.d. random variables with uniform distribution on $[-\frac{1}{2}, \frac{1}{2}]$) has the bound $p_X(t) \leq C e^{-c|t|}$. We will give two proofs of this fact. The first proof is based on properties of logarithmically concave distributions. The second proof combines the classical Bernstein-Hoeffding estimate with a simple argument based on the Fourier transform. Note that the second proof yields a somewhat stronger conclusion that $p_X(t) \leq C e^{-ct^2}$. This, in turn, improves the factor $e^{-c|L|}$ in Lemmas 4' and 4 to e^{-cL^2} .

5.1 Log-concavity

A random vector in \mathbb{R}^d having density $p(\cdot)$ is said to be log-concave if the function $\log p: \mathbb{R}^d \rightarrow \mathbb{R} \cup \{-\infty\}$ is concave. For our purposes, it suffices to know the following two basic classes of examples and a general property of log-concave densities in one dimension, given below in Lemma 12.

1. If the random vector X is uniformly distributed on a compact convex set K , then X is log-concave. Indeed, if p is the density of X , then

$$\log p(x) = \begin{cases} \log(1/\text{vol}(K)) & \text{if } x \in K, \\ -\infty & \text{if } x \notin K, \end{cases}$$

which is easily seen to be concave.

2. If X is as above (uniform on a convex set in \mathbb{R}^d), and $u \in \mathbb{R}^d$ is any fixed vector, then $\langle u, X \rangle$ is log-concave in one dimension. This is not an obvious fact, but is a consequence of the Prékopa-Leindler inequality [7, 3].

As a consequence, if V is uniformly distributed on $[0, 1]^n$ (a convex set) then, the scalar product $\langle w, V \rangle$ is a log-concave random variable in one dimension, for any $w \in \mathbb{R}^n$.

Here is the one key (and well-known to experts) property of log-concave random variables that we need.

Lemma 12. *Let X be a real-valued, symmetric, log-concave random variable with unit variance. Then the density $p(\cdot)$ of X satisfies $p(t) \leq Ce^{-c|t|}$ for all t for some constants C, c .*

The lemma remains valid if we drop the assumption of symmetry and instead assume that X has zero mean, but the proof would be a bit longer. Since we only apply this to symmetric random variables, we state only this weaker version.

Proof of Lemma 12. Since $\log p$ is concave and symmetric, it is non-increasing on $(0, \infty)$ and non-decreasing on $(-\infty, 0)$. Define $B = \inf\{t > 0: p(t) \leq \frac{1}{2}p(0)\}$. Then $0 < B < \infty$. Further, log-concavity shows that $p(kB)^{1/k}p(0)^{(k-1)/k} \leq p(B) \leq \frac{1}{2}p(0)$, whence, $p(kB) \leq p(0)2^{-k}$. By the unimodality of p , we now have the bounds

$$\begin{aligned} p(t) &\geq \frac{1}{2}p(0) \text{ if } t \in (0, B), \\ p(t) &\leq p(0)2^{-k} \text{ if } t \in (kB, (k+1)B), \quad k \geq 0. \end{aligned}$$

From these bounds, we get

$$\frac{1}{2}Bp(0) \leq \int_0^\infty p(t) dt \leq 2Bp(0), \quad (36)$$

$$\frac{1}{6}B^3p(0) \leq \int_0^\infty t^2p(t) dt \leq 12B^3p(0). \quad (37)$$

In the last inequality, we used the fact that $\sum_{k=0}^\infty 2^{-k}(k+1)^2 = 12$ (any constant upper bound would suffice). But, $\int_0^\infty p(t) dt = \frac{1}{2}$ and $\int_0^\infty t^2p(t) dt = \frac{1}{2}\text{Var}[X] = \frac{1}{2}$. Combining with (36) and (37), we see that $c \leq p(0) \leq C$ and $c \leq B \leq C$ for some numerical constants c, C .

Thus, for $t \in [-B, B]$, we have $p(t) \leq C$. Also, for $t > kB$ we have $p(t) \leq p(0)2^{-k}$ which gives the exponential upper bound $p(t) \leq Ce^{-ct}$. \square

5.2 Proof of Lemma 10

As $U = (U_1, \dots, U_n)$ is uniformly distributed on $[-\frac{1}{2}, \frac{1}{2}]^n$, we see that $X = \sum_{i=1}^n w_i U_i$ is a symmetric, log-concave random variable. Its variance is $\frac{1}{12} \sum_{i=1}^n w_i^2$. If $\sum_{i=1}^n w_i^2 = 1$, then, by Lemma 12, the density of X is bounded by $Ce^{-c|t|}$. This proves the first part of the lemma.

If $\|w\|^2 = \sum_{i=1}^n w_i^2 < 1$, then let $Y = \frac{1}{\|w\|}X$. By the first part, the density of Y is bounded by $Ke^{-\kappa|t|}$. The density of X is given by

$$p_X(t) = p_Y(t/\|w\|) \|w\|^{-1} \leq \frac{C}{\|w\|} e^{-c|t|/\|w\|}.$$

Since the function $x \rightarrow xe^{-c|t|x/2}$ is bounded by $\frac{2}{c|t|}$, we get the bound

$$p_X(t) \leq \frac{2C}{c|t|} e^{-c|t|/(2\|w\|)}.$$

For $|t| \geq 1$ and $\|w\| \leq 1$, this is less than or equal to $C'e^{-c'|t|}$. \square

5.3 Another proof of (an improved version of) Lemma 10

First, we recall the following classical inequality which goes back to Bernstein and Hoeffding.

5.3.1 The Bernstein-Hoeffding inequality

If X_1, \dots, X_n are independent random variables with zero mean, and such that $|X_i| \leq a_i$ a.s., where a_i are non-negative numbers. Then, for any $t > 0$,

$$\mathbb{P}\left\{\sum_{i=1}^n X_k \geq t\right\} \leq \exp\left\{-\frac{1}{2} \frac{t^2}{\sum_{i=1}^n a_i^2}\right\}. \quad (38)$$

5.3.2 A bound for $p_X(t)$ when $|t| \geq 1$

Now consider $X_i = w_i U_i$ as in Lemma 10. Clearly X_i are independent and $|X_i| \leq \frac{1}{2} |w_i|$. Therefore, by (38), we get

$$\mathbb{P}\{X \geq t\} \leq \exp\left\{-\frac{1}{2} \frac{t^2}{\sum_{i=1}^n (w_i/2)^2}\right\} \leq e^{-2t^2}$$

since $\sum_{i=1}^n w_i^2 \leq 1$. To get a bound on the density from the bound on the tail, we observe that X is a symmetric (i.e., $p_X(t) = p_X(-t)$) and unimodal³ (that is, p_X is non-decreasing on $(-\infty, 0]$ and non-increasing on $[0, \infty)$). Therefore, for any $t > 0$ we get

$$\frac{t}{2} p_X(t) \leq \int_{t/2}^t p_X(s) ds \leq \mathbb{P}\{X \geq t/2\} \leq e^{-t^2/2}.$$

For $t > 1$, this gives the bound $p_X(t) \leq 2e^{-t^2/2}$. By symmetry, the same bound holds for $t < -1$. This is what we set out to prove when $|t| \geq 1$. \square

5.3.3 A bound for $p_X(t)$ when $|t| \leq 1$

Here, $X = \sum_i w_i U_i$, where U_i are i.i.d. random variables uniformly distributed on $[-\frac{1}{2}, \frac{1}{2}]$, and $\sum_i w_i^2 = 1$. We will show that, for $|t| \leq 1$, $p_X(t) \leq C$.

Let δ be a small positive parameter, to be chosen near the end of the proof. If, for some i , $|w_i| \geq \delta$, then we use the estimate

$$p_X(t) \leq \frac{1}{\delta}, \quad |t| \leq 1,$$

which follows from the fact that the sup-norm of the density cannot increase under convolution. In what follows, we assume that $|w_i| < \delta$ for all $i \geq 1$.

The characteristic function of X equals

$$\mathbb{E}[e^{i\lambda X}] = \prod_{i \geq 1} \frac{\sin(\lambda w_i)}{\lambda w_i},$$

so we need to show that

$$\int_0^\infty \prod_{i \geq 1} \left| \frac{\sin(\lambda w_i)}{\lambda w_i} \right| d\lambda \leq C,$$

³ The unimodality may be proved easily by induction on the number of summands.

provided that $\sum_i w_i^2 = 1$ and that $|w_i| < \delta$. We will be using that

$$\left| \frac{\sin \xi}{\xi} \right| \leq \begin{cases} e^{-c\xi^2}, & |\xi| \leq 2, \\ \frac{1}{|\xi|}, & |\xi| \geq 1. \end{cases}$$

Let $I_j = [2^j, 2^{j+1}]$, $j \geq 1$, and let $J_k = \{i: 2^{-k} < |w_i| \leq 2^{1-k}\}$, $k \geq 1$. Then

$$\frac{1}{4} \leq \sum_{k \geq 1} |J_k| 2^{-2k} \leq 1.$$

For $\lambda \in I_j$, $i \in J_k$, we have $2^{j-k} \leq |\lambda w_i| \leq 2^{2+j-k}$, whence,

$$\left| \frac{\sin(\lambda w_i)}{\lambda w_i} \right| \leq \begin{cases} e^{-c2^{2(j-k)}}, & k \geq j, \\ 2^{k-j}, & 1 \leq k \leq j-1. \end{cases}$$

Thus,

$$\begin{aligned} \int_{I_j} \prod_{i \geq 1} \left| \frac{\sin(\lambda w_i)}{\lambda w_i} \right| d\lambda &\leq 2^j \prod_{1 \leq k \leq j-1} 2^{(k-j)|J_k|} \prod_{k \geq j} e^{-c2^{2(j-k)}|J_k|} \\ &= 2^j \exp \left[-\log 2 \sum_{1 \leq k \leq j-1} (j-k)2^{2k} 2^{-2k}|J_k| - c2^{2j} \sum_{k \geq j} 2^{-2k}|J_k| \right]. \end{aligned} \quad (39)$$

Now, we fix sufficiently large positive integers j_0 and k_0 so that, for $j \geq j_0$, we have $c2^{2j} \geq 8 \log 2$ (here, c is the constant in the exponent on the right-hand side of (39)), and that, for $j \geq j_0$, $k_0 \leq k \leq j-1$, we have $(j-k)2^{2k} \geq 8j$. Then, we take $\delta = 2^{-k_0}$. This choice guarantees that $J_k = \emptyset$ for $1 \leq k < k_0$. Then, for $j \geq j_0$,

$$\text{right-hand side of (39)} \leq 2^j \exp \left[-(8 \log 2)j \sum_{k \geq k_0} 2^{-2k}|J_k| \right] \leq 2^{-j}.$$

Therefore,

$$\int_0^\infty \prod_{i \geq 1} \left| \frac{\sin(\lambda w_i)}{\lambda w_i} \right| d\lambda = \left(\int_0^{2^{j_0}} + \int_{2^{j_0}}^\infty \right) \dots \leq 2^{j_0} + \sum_{j \geq j_0} 2^{-j} < 2^{j_0} + 1,$$

completing the proof. \square

6 Proof of Claim 11

Let V_1, \dots, V_n be i.i.d. uniform random variables, and $V_{(1)} < V_{(2)} < \dots < V_{(n)}$ be their order statistics (i.e., $V_{(j)}$ is the j th minimum of the V_i s). Since the distribution of $(V_{(1)}, V_{(2)}, \dots, V_{(n)})$ is the same as that of $((V_\sigma)_{\sigma(1)}, \dots, (V_\sigma)_{\sigma(n)})$, $\langle w, V_\sigma \rangle$ has the same distribution as $\sum_{j=1}^n w_{\sigma(j)} V_{(j)}$. Put $X_j = V_{(j)} - V_{(j-1)}$, $1 \leq j \leq n+1$, with the convention that $V_{(0)} = 0$ and $V_{(n+1)} = 1$. Thus, the lemma will follow if we consider

$$Y = \sum_{j=1}^n w_{\sigma(j)} V_{(j)} = \sum_{j=1}^n \alpha_j X_j$$

with $\alpha_j = w_{\sigma(j)} + \dots + w_{\sigma(n)}$, and show that $\text{Var}[Y] \leq \frac{1}{(n+1)(n+2)} \sum_{j=1}^n \alpha_j^2$.

To show this, we will need one property of the distribution of the vector (X_1, \dots, X_{n+1}) , that of exchangeability: if the X_i s are permuted, the resulting vector has the same distribution as (X_1, \dots, X_{n+1}) . To see this, drop one of the X_i s and write the density of the remaining n random variables with respect to Lebesgue measure on \mathbb{R}^n . A change of variables shows that the density is uniform on $\{(t_1, \dots, t_n) : t_i \geq 0, \sum_{i=1}^n t_i < 1\}$. Then, exchangeability is clear.

From exchangeability, we see that $\mathbb{E}[X_i] = \mathbb{E}[X_1]$, $\mathbb{E}[X_i^2] = \mathbb{E}[X_1^2]$ and $\mathbb{E}[X_i X_j] = \mathbb{E}[X_1 X_2]$ for all $i \neq j$. Further, $X_1 + \dots + X_n = 1$. Therefore, $\mathbb{E}[X_1] = \frac{1}{n+1}$ and

$$(n+1)\mathbb{E}[X_1^2] + n(n+1)\mathbb{E}[X_1 X_2] = 1$$

since the left hand side is $\mathbb{E}[(X_1 + \dots + X_{n+1})^2]$.

Since $X_1 = V_{(1)}$ is the minimum of n uniform random variables, we easily calculate that $\mathbb{E}[X_1^2] = \frac{2}{(n+1)(n+2)}$. Then we also get $\mathbb{E}[X_1 X_2] = \frac{1}{(n+1)(n+2)}$.

Thus, we get $\mathbb{E}[Y] = \frac{1}{n+1} \sum_{j=1}^n \alpha_j$, and

$$\begin{aligned} \mathbb{E}[Y^2] &= \frac{2}{(n+1)(n+2)} \sum_{j=1}^n \alpha_j^2 + \frac{1}{(n+1)(n+2)} \sum_{i \neq j} \alpha_i \alpha_j \\ &= \frac{1}{(n+1)(n+2)} \sum_{j=1}^n \alpha_j^2 + \frac{1}{(n+1)(n+2)} \left(\sum_{j=1}^n \alpha_j \right)^2. \end{aligned}$$

Finally,

$$\begin{aligned} \text{Var}[\langle w, V_\sigma \rangle] &= \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 \\ &= \frac{1}{(n+1)(n+2)} \sum_{j=1}^n \alpha_j^2 - \frac{1}{(n+1)^2(n+2)} \left(\sum_{j=1}^n \alpha_j \right)^2 \end{aligned}$$

which completes the proof. \square

Remark. The distribution of the X_i s appearing in the proof of Claim 11 is called the Dirichlet distribution with parameters $n+1$ and $(1, \dots, 1)$.

References

- [1] A. BLOCH, G. PÓLYA, On the roots of certain algebraic equations. Proc. London Math. Soc. **33** (1932), 102–114.
- [2] C. G. ESSEEN On the Kolmogorov-Rogozin inequality for the concentration function. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **5** (1966), 210–216.
- [3] R. J. GARDNER, The Brunn-Minkowski inequality.
<http://faculty.wvu.edu/gardner/gorizia12.pdf>
- [4] M. KAC, Signal and noise problems. Amer. Math. Monthly **61**, (1954), 23–26.
- [5] J.P. KAHANE, Some random series of functions, second ed., Cambridge University Press, Cambridge, 1985.
- [6] G. PÓLYA, G. SZEGŐ, Problems and Theorems in Analysis II. Springer-Verlag, Berlin, 1998.
- [7] A. PRÉKOPA On logarithmic concave measures and functions. Acta Sci. Math. (Szeged) **34** (1973), 335–343.
- [8] K. SÖZE, Real zeroes of random polynomials, I. Flip-invariance, Turán’s lemma, and the Newton-Hadamard polygon.